THE QUARTERLY JOURNAL OF EXPERIMENTAL PSYCHOLOGY

Volume 21 1969

Editorial Board
R. J. Audley
A. D. Baddeley
G. S. Brindley
D. E. Broadbent
J. Brown
V. R. Cane
R. Davis
R. L. Gregory
I. M. L. Hunter
H. Kay
N. Moray
P. M. A. Rabbitt
R. L. Reid
N. S. Sutherland
M. Treisman
A. M. Uttley
P. H. Venables
D. M. Vowles
L. Weiskrantz
P. Whittle

Editor
A. J. WATSON

Assistant Editor
M. P. HAGGARD

PUBLISHED FOR THE EXPERIMENTAL PSYCHOLOGY SOCIETY
BY ACADEMIC PRESS LONDON AND NEW YORK

Copyright © 1969 by Experimental Psychology Society

ALL RIGHTS RESERVED

No part of this volume may be reproduced in any form, by photostat, microfilm, or any other means, without written permission from the publishers.

Printed in Great Britain

Volume 21

NUMBER 1, FEBRUARY 1969

JOHNSON-LAIRD, P. N., On understanding logically complex sentences . 1
WASON, P. C. and JOHNSON-LAIRD, P. N., Proving a disjunctive rule . 14
SANDERS, A. F. and SCHROOTS, J. J. F., Cognitive categories and memory span. III. Effects of similarity on recall 21
MARKOVA, I., Hypothesis formation and problem complexity . . . 29
SAMPSON, H., Recall of digits projected to temporal and nasal hemi-retinas 39
HABER, R. N. and STANDING, L. G., Direct measures of short-term visual storage 43
HUGHES, R. N. and BLACKMAN, D. E., Stimulus change effects of forcing on maze arm selection in hamsters 55
DOTY, B. A. and COMBS, W. C., Reversal learning of object and positional discriminations by mink, ferrets and skunks 58
DUTCH, J., Visual complexity and stimulus pacing in chicks . . . 63
KINSBOURNE, M. and COUGHLIN, P. J., The demonstration of dichoptic flicker effects by use of Polaroid 67
BOWMAN, R. and SUTHERLAND, N. S., Discrimination of "W" and "V" shapes by goldfish 69
SUTHERLAND, N. S. and WILLIAMS, C., Discrimination of checkerboard patterns by rats 77
NORMAN, D. A., Memory while shadowing 85
BOOK REVIEWS 94
PROCEEDINGS OF THE EXPERIMENTAL PSYCHOLOGY SOCIETY, 1968 . . i
MEMBERS OF THE EXPERIMENTAL PSYCHOLOGY SOCIETY iv

NUMBER 2, MAY 1969

HINDE, R. A., Control of movement patterns in animals 105
DODWELL, P. C., STANDING, L. G. and THIO, H., Are thresholds reduced by illusions? An attempt at replication 127
BROWN, H. V. and MILES, T. R., Prior stimulation and the perception of causality 134
RYAN, J., Grouping and short-term memory: different means and patterns of grouping 137
RYAN, J., Temporal grouping, rehearsal and short-term memory . . 148
RASHOTTE, M. E. and SURRIDGE, C. T., Partial reinforcement and partial delay of reinforcement effects with 72-hour intertrial intervals and interpolated continuous reinforcement 156
OATLEY, K. and TONGE, D. A., The effect of hunger on water intake in rats 162

OATLEY, K., ROBERTSON, A. and SCANLAN, P. M., Judging the order of visual stimuli 172
GILSON, E. Q. and BADDELEY, A. D., Tactile short-term memory . . 180
BRIGGS, G. E. and SWANSON, J. M., Retrieval time as a function of memory ensemble size 185
BOOK REVIEWS 192

NUMBER 3, AUGUST 1969

GERVER, D., Effects of grammaticalness, presentation rate and message length on auditory short-term memory 203
SHAFFER, L. H. and HARDWICK, J., Errors and error detection in typing 209
FREUND, R. D., BRELSFORD, J. W., JR. and ATKINSON, R. C., Recognition vs. recall: storage or retrieval differences 214
HUMPHREY, N. K. and WEISKRANTZ, L., Size constancy in monkeys with inferotemporal lesions 225
GOGEL, W. C., The effect of object familiarity on the perception of size and distance 239
ANDERSON, N. S., The influence of acoustic similarity on serial recall of letter sequences 248
HANDEL S. and BUFFARDI, L., Using several modalities to perceive one temporal pattern 256
HUPPERT, F. A. and DEUTSCH, J. A., Improvement in memory with time 267
COHEN, G., Some evidence for parallel comparisons in a letter recognition task 272
ABRAMS, K. and BEVER, T. G., Syntactic structure modifies attention during speech perception and recognition 280
BOOK REVIEWS 291

NUMBER 4, NOVEMBER 1969

BURNS, B. D., MANDL, G., PRITCHARD, R. and WEBB, C., The perception of briefly exposed point-sources of light 299
SEYMOUR, P. H. K., Matching latencies for world-shape pairs . . 312
BIEDERMAN, G. B., The role of the CS and avoidance period duration in discrimination avoidance conditioning 322
SUTHERLAND, N. S. and BOWMAN, R. Discrimination of circles and squares with and without knobs by goldfish 330
STILL, A. W., Proactive interference and spontaneous alternation in rats . 339
ROSS, H. E., When is a weight not illusory? 346
FISHER, G. H., An experimental study of angular subtension . . . 356
WILTON, R. N., STRONGMAN, K. T. and NERENBERG, A. Some effects of frustration in a free responding operant situation 367
SHAFFER, L. H. and HARDWICK, J. Reading and Typing 381
BOOK REVIEWS 384
INDEX TO AUTHORS 387

Q. Jl exp. Psychol. (1969) **21**, 1–13

ON UNDERSTANDING LOGICALLY COMPLEX SENTENCES

P. N. JOHNSON-LAIRD

Department of Psychology, University College London

An experimental investigation was made into the meaning of eight types of *doubly-quantified* sentence, e.g. "*Every* medicine cures *some* disease," "*Some* disease is cured by *every* medicine." All the sentences were ambiguous, depending upon the interpretation of the quantifiers. Subjects classified diagrams representing different specific situations as truthfully or falsely described by the sentences. The classifications revealed that the order of occurrence of the two quantifiers had a crucial effect, causing active and correlative passive to receive different interpretations. This suggested that in the process of understanding an ambiguous sentence a bias towards one intepretation may be created by word order.

Introduction

There is a controversy about whether the active and passive voice are *always* synonymous (cf. Chomsky, 1957, 1965; Katz and Postal, 1964; Ziff, 1966; Katz and Martin, 1967). It seems to have arisen, in part, because of a failure to distinguish probable interpretations from possible interpretations. Consider, for example, this pair of *quantified* sentences:

(1) *All* philosophers have read *some* books.

(2) *Some* books have been read by *all* philosophers.

Both sentences, isolated from any context, are ambiguous in the same way. They mean either (a) all philosophers have read some books *or other*, or (b) all philosophers have read some books *in particular*. But, although both sentences have the same possible interpretations, it seems intuitively that these interpretations are not equally probable—they are not equally privileged. Sentence (1) seems more likely to receive interpretation (a), and sentence (2) seems more likely to receive interpretation (b). One aim of the present experiment was to test whether there was this difference in the privileged interpretations of a variety of logically complex sentences.

Is there any explanation why these sentences should have different privileged interpretations? A preliminary answer is provided by considering their symbolic form within the quantificational calculus (cf. Ch. 4 of Suppes, 1957). The existential quantifier (Ex) stands for "there exists at least one x . . .," and the universal quantifier (x) stands for "for any x. . . ." Let x range over philosophers,

y range over books, and R stand for the relation of "reading," then it is possible to express the two interpretations of the sentences in the following form:

(a) $(x)(Ey)(xRy)$ Every philosopher has read some books *or other*.

(b) $(Ey)(x)(xRy)$ There are some books *in particular* that every philosopher has read.

These symbolic interpretations have their quantifiers in the same order as the actual sentences for which they are assumed to be privileged. Hence, an explanation for the effect is that word order determines which interpretation is privileged. What comes first in a sentence receives more emphasis than what comes later (cf. Johnson-Laird, 1968*a*, *b*) and the first quantifier will, in the technical sense, *bind* the second quantifier. This is not a complete explanation, however, until an analogue for the technical notion of binding is found in everyday language. Such an analogue is revealed by an examination of the expressions which realize the existential quantifier, e.g. "some," "at least one," "several," "a few," etc. All of these expressions are essentially ambiguous. Compare the probable interpretations of "John has several girl friends" and "John always has several girl friends"; in the first sentence the reference is to some girls in particular, but in the second sentence it is to some girls or other. Where there is only a single quantifier in the sentence the ambiguity is trivial, but where there are two or more quantifiers it may be crucial. When "some" occurs first it will receive a greater emphasis according to the hypothesis about word order. This emphasis is more in accord with "some in particular" than "some or other," since emphasis is given to particulars rather than to vague generalities. The converse argument applies where the existential quantifier comes after the universal quantifier. This explains the privileged interpretations of sentences (1) and (2). But the explanation also applies to other sentences where intuition is a less certain guide, and by examining these sentences experimentally the theory is put to a more stringent test.

TABLE I

Twelve basic quantified "two-place" relations

1.	$(x)\ (Ey)\ (xKy)$	Every man knows some woman (or other).
2.	$(Ey)\ (x)\ (xKy)$	Some woman (in particular) is known by every man.
3.	$(y)\ (Ex)\ (xKy)$	Every woman is known by some man (or other).
4.	$(Ex)\ (y)\ (xKy)$	Some man (in particular) knows every woman.
5.	$(x)\ (Ey) - (xKy)$	Any man does not know some woman (or other).
6.	$(Ey)\ (x) - (xKy)$	Some woman (in particular) is not known by any man.
7.	$(y)\ (Ex) - (xKy)$	Any woman is not known by some man (or other).
8.	$(Ex)\ (y) - (xKy)$	Some man (in particular) does not know any woman.
9.	$(x)\ (y)\ (xKy)$	Every man knows every woman.
10.	$(x) - (Ey)\ (xKy)$	No woman is known by any man whatsoever.
11.	$(Ex)\ (Ey)\ (xKy)$	Some man knows some woman.
12.	$(Ey) - (x)\ (xKy)$	Some woman (in particular) is not known by every man.

(Ej) denotes the existential quantifier, (j) denotes the universal quantifier, and "—" denotes negation. It is assumed that x ranges over men, y ranges over women, and K denotes the relation of "knowing".

There are 12 basic quantified "two-place" relations, and these are listed in Table I in a symbolicm for and in a more or less idiomatic translation ("—" is the sign for negation). There are other logically equivalent formulations of these relations, any one of which may be expressed in English in a variety of ways. But the material selected for investigation was of the same form as sentences (1) to (8). This material was chosen in order to investigate three variables: *polarity*—whether the sentence is affirmative or negative, *voice*—whether the sentence is active or passive, and *word order*—whether "some" is in the grammatical (surface) subject or object of the sentence.

There is another reason for investigating these sentences. A traditional problem in the psychology of reasoning is to determine the influence of linguistic form upon judgements of validity (cf. Woodworth and Sells, 1935; Sells, 1936). The experiments typically utilize syllogisms with singly-quantified sentences; but inferences may also be made with doubly-quantified sentences and these, too, may be influenced by linguistic form. Experiments could utilize either the unambiguous but stilted jargon of the logician ("For any man, there exists a woman such that the man knows the woman") or the idiomatic but ambiguous sentences of everyday discourse ("Every man knows some woman"). What, in fact, is needed is a preliminary study of the idiomatic material, since its very ambiguity may be a significant factor in reasoning. The present experiment was intended as a preliminary to a further study of inference.

Method

Task

Each subject was presented with a series of sentences. For each one, ten diagrams—representing different states of affairs—had to be classified as truthfully or falsely described by the sentence. The diagrams were chosen so that their classification would elucidate the subjects' intepretations of the sentences.

Design

Each subject was his own control and received the eight different forms of sentence. In order to reduce residual effects from one item to the next, each of the sentences had a different lexical content. Such effects seemed likely to be greater for logical form than for lexical content, hence, as a procedural variable, a series of 24 counterbalanced orders of presentation was derived from three 8 × 8 Williams' squares (cf. Edwards, 1963, p. 275). A separate order for the lexical material was assigned to each Williams' square; this provides only a crude control for effects associated with the lexical material, but it does provide an adequate control for any interaction between logical form and lexical content. The order of presentation of the diagrams was randomly determined for every trial.

Materials

The eight logical forms were those of sentences (1) to (8) except that the words in parentheses were omitted. The universal quantifier differs between the affirmative and negative sentences. In a pilot study, the use of "any" gave rise to a peculiar difficulty especially with the affirmative sentences. Subjects tended to regard a sentence like "Any medicine cures some disease" as true if a single medicine, chosen at random by the subject, cured some disease. Thus, the truth of the sentence depended upon "the luck of the draw." "Any" was accordingly used only where there was no alternative, that is, in the negative sentences, since "Some medicine does not cure *any* disease" differs in meaning to "Some medicine does not cure *every* disease."

The lexical materials are stated below in the same logical form throughout:

(a) Every man knows some woman.
(b) Every car overtakes some bus.
(c) Every job requires some skill.
(d) Every child loves some toy.
(e) Every brother hates some sister.
(f) Every liquid dissolves some solid.
(g) Every medicine cures some disease.
(h) Every animal possesses some instinct.

These materials were selected so as to vary the types of noun (e.g. human, abstract, etc.) and the reversibility of the sentences, i.e. whether the two nouns could be interchanged without producing a semantic anomaly (cf. Slobin, 1966).

The diagrams consisted of 3 × 3 matrices with each cell containing either a "1" or a "0": a typical matrix is shown in Figure 1. The ten matrices are shown in Table II. To understand the way the matrices have to be interpreted, imagine that they are paired with the sentence "Some woman is known by every man." Each row in the matrices then represents a man, and each column represents a woman. A "1" indicates that the man in whose row it is *knows* the woman in whose column it is. A "0" indicates that the man does *not* know the woman. Thus, the situation depicted in, say, matrix (8) is likely to be classified as truthfully described by the above sentence.

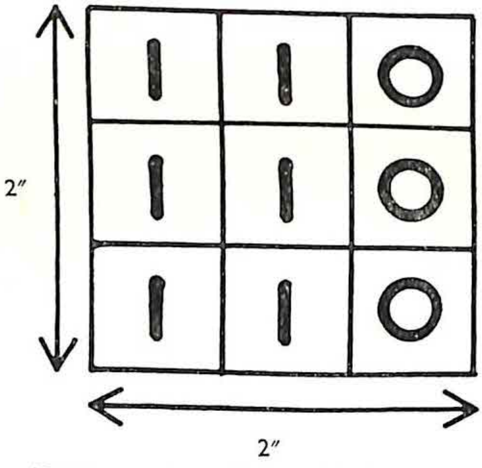


FIGURE 1. A typical matrix diagram.

The choice of the particular ten matrices was based on two grounds. First, they were chosen as a representative sample from the complete set of such matrices, ignoring rotations or other transformations which had no effect upon how the matrices ought to be classified. Second, they were chosen so that the likely interpretations of sentences would be clearly reflected in the *pattern* of the subjects' choices. Three pairs of matrices, (3) and (8), (4) and (7), and (5) and (9), were chosen so that the logical classification would be the same for both members of the pair. These mutually redundant pairs provide a simple test of whether an interpretation is within the quantificational calculus.

The rows of the matrices always represented the logical (deep structure) subject and the columns the logical object of the sentences. This was likely to enhance any effect due to voice, but it was nevertheless left uncontrolled so as to minimize the difficulty of the task. To help subjects to identify rows and columns, a large "backing" card (9 in. × 6 in.) on which there was an appropriately labelled blank matrix was on view throughout the classification. Eight backing cards were prepared: one for each set of lexical material, with the labels on the rows and columns typed in capitals.

Each of the 64 sentences (8 logical forms × 8 lexical contents) was typed on a plain 5 in × 3 in. card. The matrices, 2 in. × 2 in., were drawn in heavy ink at one end of a similar card, so that the subject could hold a card in his hand and clearly see the matrix.

Predictions

The nature of the predictions was slightly unusual in so far as there were two levels within them. Table II gives what will be termed the *privileged* classifications of the matrices. They rest on the assumption that subjects will give each sentence its privileged interpretation and

TABLE II

Ten matrix diagrams and the privileged classifications for the basic quantified relations

	1	2	3	4	5	6	7	8	9	10
	000 000 000	100 000 000	100 100 100	100 010 001	111 000 000	111 100 100	110 101 011	110 110 110	111 111 000	111 111 111
1. Every man knows some woman	0	0	1	1	0	1	1	1	0	1
2. Some woman is known by every man	0	0	1	0	0	1	0	1	0	1
3. Every woman is known by some man	0	0	0	1	1	1	1	0	1	1
4. Some man knows every woman	0	0	0	0	1	1	0	0	1	1
5. Any man does not know some woman	1	1	1	1	0	0	1	1	0	0
6. Some woman is not known by any man	1	1	1	0	0	0	0	1	0	0
7. Any woman is not known by some man	1	1	0	1	1	0	1	0	1	0
8. Some man does not know any woman	1	1	0	0	1	0	0	0	1	0
9. Every man knows every woman	0	0	0	0	0	0	0	0	0	1
10. No woman is known by any man	1	0	0	0	0	0	0	0	0	0
11. Some man knows some woman	0	1	1	1	1	1	1	1	1	1
12. Some woman is not known by every man	1	1	1	1	1	1	1	1	1	0

Note: Each row of a matrix represents a man, and each column represents a woman. A "1" under a matrix indicates that the matrix is classified as truthfully described, a "0" indicates that it is classified as falsely described. For convenience, the privileged classifications for sentences 9 to 12, not actually used in the experiment, have been stated.

then classify the matrices in a strictly logical fashion. These predicted patterns of classification may be summarized in the following way:

Sentences	*Matrices truthfully described*	*Matrices falsely described*
(1)	At least one “1” in every row.	At least one row of “o's.”
(2)	At least one column of “1's.”	At least one “o” in every column.
(3)	At least one “1” in every column.	At least one column of “o's,”
(4)	At least one row of “1's,”	At least one “o” in every row.
(5)	At least one “o” in every row.	At least one row of “1's”.
(6)	At least one column of “o's”.	At least one “1” in every column.
(7)	At least one “o” in every column.	At least one column of “1's.”
(8)	At least one row of “o's.”	At least one “1” in every row.

It is clear that subjects can depart from the privileged classifications in a number of ways: they may produce the unprivileged classification (i.e. the one appropriate to the correlated sentence of the opposite voice), or a classification which is neither privileged nor unprivileged although still within the quantificational calculus (some Boolean function of the patterns in Table II), or finally a classification which is outside the quantificational calculus altogether.

Although subjects were expected to produce the privileged classifications, they were not expected to perform in every respect like trained logicians. Hence, a series of lower level predictions were made concerning departures from the privileged classifications: certain aspects of the sentences would cause difficulty, i.e. longer classification times and departures from the privileged classifications. It was predicted (i) that negatives would cause more difficulty than affirmatives. Negatives customarily cause difficulty (Wason, 1959, 1961) unless they occur in plausible contexts (Wason, 1965). It was also predicted (ii) that passives would cause more difficulty than actives, although this would be less than the effect caused by negatives (cf. McMahon, 1963; Gough, 1965; Slobin, 1966). Finally, it was predicted (iii) that the word order of the quantifiers would have an effect. The privileged interpretation should be easier to make when “some” is in the grammatical subject than when it is in the grammatical object. The argument underlying this prediction is that when “some” occurs in the object, it is taken as “some or other” almost by default, i.e. “some” is *not* in the subject therefore it does not mean “some in particular”. Accordingly, this should lead to a greater difficulty in classifying the matrices.

Subjects

The subjects were 24 first-year undergraduates (14 male, 10 female) in the Department of Psychology, University College London. They were native-speakers of English but they were not familiar with the quantificational calculus.

Procedure

The subjects were allocated to a particular combination of logical and lexical material, and tested individually.

The subject sat at the end of a table which was screened off from the experimenter's part of the table. The experimenter read the instructions to the subject. These explained that the task was to determine what situations were truthfully (and falsely) described by a series of sentences. Each situation would be depicted by a diagram. A matrix was then introduced and the way it was to be interpreted was described in detail. The classification procedure was explained, and the subject was told that he could not ask any questions after the backing card for a sentence had been described. The subject was also told that he would be timed but that this must not disturb him: “I don't want you to race blindly through the task. Just imagine that somebody else has said the sentence to you, make up your mind what you think it means, and then make your classification. Make sure that your classification truly represents your considered opinion about the meaning of the sentence.” Finally, the subject was warned not to let his judgements be influenced by whether he thought the sentences were *really* true (or false).

The experimenter produced the appropriate backing card and the matrices, and described the lexical content of the sentence. The stopwatch was started as soon as the sentence was placed in front of the subject; it was stopped as soon as the last card was classified. The subjects were allowed to change their classifications, and the times recorded include intervals taken up by this. Between trials the experimenter recorded the time and the classifications, and re-ordered the matrices for the next trial, all of which took up an inter-trial interval of about 25 sec.

Results

Classifications

The proportion of subjects making the privileged, unprivileged, and other classifications for each form of sentence is shown in Figure 2. There was no reliable difference in performance with the different lexical materials and no interaction between them and the logical forms, hence the analysis was on pooled data. The modal classification for each of the eight logical types was the predicted privileged one. The average number of privileged classifications per subject was 4·7 (out of a maximum of 8); since the probability of a given classification by chance is less than one in a thousand, it is evident that there was a very reliable bias in favour of the predicted classifications.

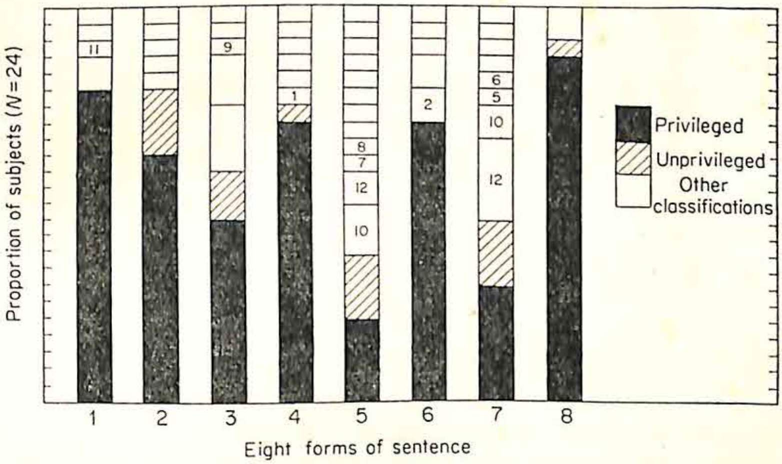


FIGURE 2. The proportions of subjects making the privileged, unprivileged, and other classifications for each form of sentence. (A number in a column indicates that the classification was privileged for the corresponding sentence e.g. 3 subjects gave sentence 5 the privileged interpretation of sentence 10.)

Since a major concern was whether the classifications were privileged or unprivileged, each classification was scored with respect to both of these interpretations. The higher score was taken as an index of which interpretation had been made. This made it possible to take all the data into account and to assign to each subject a number, ranging from +8 to −8, as a function of the degree to which he made the privileged interpretations. There was an overall mean of +5·2, and all 24 subjects had a positive score, which is highly significant on a sign test ($P = 0{\cdot}5^{24}$).

Departures from the privileged classifications

The maximum number of departures from the privileged classification which could be made on any trial was ten, since there were ten matrices to be classified. The mean number of such departures for each of the eight forms of sentence is shown in Table III, along with the classification times. An analysis of variance ($2 \times 2 \times 2$ fixed factors with repeated measures) indicated that both polarity and word order produced reliable differences in the predicted direction, with $F_{1,161}$ ratios of 11·2 ($P < 0·01$) and 16·5 ($P < 0·001$), respectively. However, there was also a significant interaction between these two factors ($F_{1,161} = 19·1$, $P < 0·001$). Neither voice nor any of the remaining interactions yielded significant effects. Departures from the privileged classifications showed no signs of diminishing with practice.

TABLE III

Mean number of departures from the privileged classifications and mean classification times for the eight forms of sentence

	Affirmative				Negative				
	Subject		Object		Subject		Object		
Sentence no.	Active (4)	Passive (2)	Active (1)	Passive (3)	Active (8)	Passive (6)	Active (5)	Passive (7)	Overall
Mean departures from privileged classifications	0·8	0·9	0·4	1·1	0·3	0·9	2·4	2·3	1·1
Mean classification times (sec.)	40·2	44·6	44·5	52·4	50·7	48·0	64·5	74·4	52·4

Note: the word order factor, indicated by "subject" and "object", refers to whether "some" was in the grammatical subject or object of the sentence. The maximum possible number of departures from the privileged classification on any trial was 10. The sentence numbers quoted are those used in Table I.

If the subjects had performed as predicted in Table II, they would have classified an average of 5 out of 10 matrices as true on each trial. In fact, the overall mean was 4·5 matrices classified as true: 18 subjects had averages less than 5, three subjects had averages of greater than 5, and the remaining three subjects had averages of exactly 5. Hence, there was a significant bias ($P < 0·002$ on a sign test) in favour of classifying a matrix as falsely described by the sentence. Performance did not differ reliably between matrices: they all tended to be classified as true too few times.

Classification times

The mean classification times, from the moment the sentence was presented until the classification was complete, are shown for the eight forms of sentence in Table III. An analysis of variance indicated that there were significant effects due to polarity ($F_{1,161} = 9·2$, $P < 0·01$) and word order ($F_{1,161} = 8·1$, $P < 0·01$). Neither voice nor any interactions yielded reliable effects, though Table III indicates a trend in the direction of an interaction between polarity and word order. There was a marked learning effect in the classification times of all eight forms of sentence.

Qualitative results

Relatively few classifications were outside the quantificational calculus, i.e. involved an inconsistent classification of a mutually redundant pair of matrices. Sixteen subjects were responsible for a total of 31 such classifications (16·1 per cent of all classifications). This percentage, however, constitutes just under half of those interpretations which were neither privileged nor unprivileged. It was often the case that a plausible account could be given for the interpretations outside the quantificational calculus. For example, three subjects presented with sentence (3) of the form "Every woman is known by some man" produced the privileged classification except that they classified matrix (9) as falsely described by the sentence. It seems they had taken the some *or other* interpretation so much to heart that they refused to countenance the regularity of two particular men knowing all the women. Conversely, with sentence (4) of the form "Some man knows every woman," one subject took the some *in particular* interpretation so strictly that his classification represents one particular man knowing every woman and no other man knowing any women.

The interpretations which were within the quantificational calculus but neither privileged nor unprivileged are harder to explain. Some of them seem to be no more than mistakes, e.g. (1) "Every man knows some woman" being interpreted as (11) "Some man knows some woman." Others, particularly those involving Boolean combinations of the basic sentences, seem to have some obscure motivation, e.g. four subjects interpreted (3) "Every woman is known by some man" to imply in addition that (1) "Every man knows some woman." This is, perhaps, a case of subjects' judgements being influenced by the real world. Figure 2 shows that sentences (5) and (7) tended to receive a variety of alternative interpretations; as we shall see, these were *not* primarily due to misinterpretations of the sentences.

Subjects rarely interpreted "some" to imply "not all." Thus, for example, "Some medicine cures every disease" was deemed quite consistent with the situation in which every medicine cures every disease. This *logical* performance is in contrast to the view of Chapman and Chapman (1959) that, with singly quantified sentences, "Some A is B" is usually taken to imply "Some A is not B." Yet in ordinary discourse "some" often does imply "not all"; thus it may be that this implication is less strong with doubly quantified sentences. A related finding was that "some N"—where N stands for a singular noun—was rarely interpreted to mean just a single entity. One such case has been mentioned, there were only two others.

Introspective reports

No subject spontaneously complained about the ambiguity of the sentences, though it was evident that at least ten subjects had noticed the ambiguity of "some." A number of subjects were able to state the features of the sentences which had caused them difficulty. These reports varied from general statements referring to "the negative sentences" or "the sentences with 'any' and 'some' " to the more sophisticated reports of the five subjects who said that "some" could be either singular or plural. This distinction was made even when it was clear that the subject had never interpreted "some" as singular. It represents perhaps the nearest that subjects could get to stating the distinction between some *in particular* and some *or other*.

Only two subjects mentioned the passive voice; one subject said that he was surprised because some of the sentences could have been more simply expressed in the active voice, the other subject had "switched the sentences round when they were passives" though, in fact, this subject did *not* interpret the passives as synonymous with their active correlates. There were just seven occasions (out of 96) on which an active and correlative passive received the same interpretation.

Discussion

The task which subjects had to perform in the present experiment was difficult. It involved interpreting logically complex sentences with respect to a series of rather abstract situations. Nevertheless, the subjects consistently made the privileged classifications, with active and correlative passive sentences receiving different interpretations. The departures from the privileged classifications indicate that, as predicted, there was more of a consensus about the meaning of affirmative than negative sentences. There was, again as predicted, more of a consensus when "some" was in the grammatical subject than when it was in the grammatical object. Unfortunately for the assumptions on which these predictions were based, there was an unexpected interaction between the two factors: word order had an effect on the departures from the privileged classifications only in the case of negative sentences (cf. Table III).

The main question which the interaction raises concerns sentences (5) and (7), e.g.:

(5) Any child does not love some toy.

(7) Any toy is not loved by some child.

Why did these sentences receive more diverse interpretations than any others including those to which they are related by voice? One possibility is that these sentences are slightly anomalous, and one might even suppose that they are meaningless. Yet the answer, suggested by the subjects' actual classifications, is that they have more *possible* interpretations. This extra and unforeseen ambiguity is due to two factors.

First, the word "some" in a predicate with a negative verb-phrase has an additional ambiguity. So far, we have been considering the word pronounced [sʌm] which tends to occur with count nouns. There is another word, pronounced [sm̩], which tends to occur with mass nouns (cf. Bolinger, 1960). The difference between these two words is important in negative sentences: compare "John doesn't like some ([sʌm]) people" with "John doesn't want some ([sm̩]) butter." The latter is logically equivalent to "John wants no butter whatsoever" but the former is not equivalent to "John likes no people whatsoever." A written sentence like

"John doesn't want some pens" may thus receive any of three interpretations: John wants no pens whatsoever, John doesn't want some particular pens, or John doesn't want some pens or other (i.e. he doesn't want just any or every pen).

Second, sentences with a negative verb-phrase and a realization of the universal quantifier in their subject are ambiguous. For example, "All women do not lie" may mean either that lies are not told by *any* woman (no woman lies) or that lies are not told by *every* woman (not all women lie). This ambiguity about the scope of negation occurs whether the verb is transitive or intransitive, whether it is active or passive, and whether the quantifier is "every," "any" or "all." It even occurs in attributive descriptions like "All the square is not red" (Wason, personal communication).

Combining the two sources of ambiguity yields six possible interpretations for sentences (5) and (7). They may both receive any of the following interpretations, logically equivalent to those predicted for sentences (5), (6), (7), (8), (10) and (12):

(5) Some toy or other is not loved by any child.

(6) Some toy in particular is not loved by any child.

(10) No toy whatsoever is loved by any child.

(8) Some toy or other is not loved by every child.

(12) Some toy in particular is not loved by every child.

(7) No toy whatsoever is loved by every child.

It is a moot point whether these interpretations are possible for the voice-correlated sentences (6) and (8):

(6) Some toy is not loved by any child.

(8) Some child does not love any toy.

On balance, it would seem that they are possible: "Some ([sm̩]) child *isn't* sick" may be interpreted as "No child is sick," and "John isn't loved by any child" may be interpreted as "John isn't loved by *just* any (every) child." Hence, theoretically it would seem that sentences (5), (6), (7) and (8) are all ambiguous in the same way.

This startling consequence is not borne out in actual performance. Sentences (6) and (8) are interpreted very much as though they were unaffected by the extra sources of ambiguity and are clearly differentiated one from the other. Word order is crucial: "Some" at the beginning of a negative sentence is even less likely to elicit the interpretation "none" than the interpretation "some or other"; "any" at the end of a negative sentence is more likely to elicit the interpretation "none" than the interpretation "not all." Figure 2 shows that the most likely interpretations for sentences (5) and (7) are their respective privileged and unprivileged interpretations, though the other possibilities do all occur. This suggests that sentences like "Any man does not know some woman" tend to receive interpretations consistent with "Some woman is not known by *any* man," whereas "Any woman is not known by some man" tends to receive interpretations consistent with "Some woman is not known by *every* man." This is an effect on meaning which occurs in performance and which is due to voice and *not* to word-order changes brought about by voice (cf.

Johnson-Laird, 1968*b*). It seems that the scope of negation tends to be minimized with respect to the "actor," but maximized with respect to the "acted-upon."

There are a number of possible explanations for the differences in the classification times. It is possible that they are due to the pattern recognition aspects of the task, although the majority of subjects adopted no such explicit strategy. An informal study in which four subjects classified the matrices according to specified *patterns* showed that the average times were 8·0 sec. for "at least one column of "0's"" and 9·4 sec. for "at least one '1' in every row." The average times for the sentences were all over 40 sec. It therefore seems unlikely that the pattern recognition aspects of the task were important. Since there was no reliable interaction between polarity and word order, it is possible that these variables had their effects on the classification times for the predicted reasons. There is, however, a strong likelihood that it was the ambiguity of the sentences which was crucial. Recent studies (MacKay, 1966; MacKay and Bever, 1967) have indicated that the more ambiguous the sentence or phrase, the longer it takes to interpret or to complete.

It seems that the negation of falsity is not psychologically equivalent to truth. The reliable bias towards classifying matrices as falsely described indicates that when subjects were in doubt, they chose falsity as the "safer" classification. When neither the privileged nor the unprivileged classification was made, there was a tendency to go outside the quantificational calculus, e.g. one and only one man does not know any woman. But, often this could be explained as a zealous, if naïve, attempt to distinguish some *or other* from some *in particular*.

The fact that both privileged and unprivileged interpretations occurred demonstrates both surface and deep-structure processing of the sentences. Clark and Begun (1968) have shown both levels of processing operating in tasks where subjects had to detect or correct semantic anomalies. Katz and Postal (1964) proposed that, for models of competence, all semantic interpretation is performed on deep structure, i.e. grammatical transformations do not change meaning. It might therefore be supposed that, in actual performance, the listener needs only to extract the deep structure in order to understand the sentence. The assignment of surface structure would occur only in so far as it was necessary for the assignment of deep structure and, more importantly, it would be held only temporarily in immediate memory. This is consistent with the fact that individuals remember sense not syntax. But it is this view which the present experiment casts some doubt upon: left-to-right processing *is* important because it biases the interpretative machinery. Surface structure provides clues to deep structure, but these clues—and in particular word order—may point to only one interpretation of an ambiguous sentence. There is one difficulty with this argument, due to the lack of an account of the semantics of quantifiers. It is possible that the particular ways in which quantifiers are introduced into deep structures, or changed when deep structures are transformed into surface structures, would yield radical differences in the deep structures of voice-correlated sentences. It is hoped that this study of performance has demonstrated the need for a linguistic account of the semantics of quantifiers.

The author is very much indebted to Dr. P. C. Wason for his inestimable advice and encouragement, and for a critical reading of an earlier version of this paper.

References

BOLINGER, D. (1960). Linguistic science and linguistic engineering. *Word*, **16**, 374–91.

CHAPMAN, L. J. and CHAPMAN, J. P. (1959). Atmosphere effect re-examined. *J. exp. Pyschol.* **58**, 220–6.

CLARK, H. and BEGUN, J. S. (1968). The use of syntax in understanding sentences. *Br. J. Psychol.* **59**, 219–29.

CHOMSKY, N. (1957). *Syntactic Structures.* The Hague: Mouton.

CHOMSKY, N. (1965). *Aspects of the Theory of Syntax.* Cambridge, Mass.: M.I.T.

EDWARDS, A. E. (1963). *Experimental Design in Psychological Research.* Revised edition. London: Holt, Rinehart and Winston.

GOUGH, P. B. (1965). Grammatical transformations and speed of understanding. *J. verb. Learn. verb. Behav.* **4**, 107–11.

JOHNSON-LAIRD, P. N. (1968*a*). The interpretation of the passive voice. *Q. Jl exp. Psychol.* **20**, 69–73.

JOHNSON-LAIRD, P. N. (1968*b*). The choice of the passive voice in a communicative task. *Br. J. Psychol.* **59**, 7–15.

KATZ, J. J. and MARTIN, E. (1967). The synonymy of actives and passives. *Philos. Rev.* **76**, 476–91.

KATZ, J. J. and POSTAL, P. M. (1964). *An Integrated Theory of Linguistic Descriptions.* Cambridge, Mass.: M.I.T.

MACKAY, D. G. (1966). To end ambiguous sentences. *Percept. Psychophys.* **1**, 426–36.

MACKAY, D. G. and BEVER, T. G. (1967). In search of ambiguity. *Percept. Psychophys.* **2**, 193–200.

MCMAHON, L. E. (1963). *Grammatical Analysis as Part of Understanding a Sentence.* Unpublished doctoral thesis, Harvard University.

SELLS, S. B. (1936). The atmosphere effect: an experimental study of reasoning. *Arch. Psychol.* **29**, 3–72.

SLOBIN, D. I. (1966). Grammatical transformations and sentence comprehension in childhood and adulthood. *J. verb. Learn. verb. Behav.* **5**, 219–27.

SUPPES, P. (1957). *Introduction to Logic.* London: Van Nostrand.

WASON, P. C. (1959). Processing of positive and negative information. *Q. Jl exp. Psychol.* **11**, 92–107.

WASON, P. C. (1961). Response to affirmative and negative binary statements. *Br. J. Psychol.* **52**, 133–42.

WASON, P. C. (1965). The contexts of plausible denial. *J. verb. Learn. verb. Behav.* **4**, 7–11.

WOODWORTH, R. S. and SELLS, S. B. (1935). An atmosphere effect in formal syllogistic reasoning. *J. exp. Psychol.* **18**, 451–60.

ZIFF, P. (1966). The nonsynonymy of active and passive sentences. *Philos. Rev.* **75**, 226–32.

Received 10 *May* 1968

Q. Jl exp. Psychol. (1969) **21**, 14–20

PROVING A DISJUNCTIVE RULE

P. C. WASON AND P. N. JOHNSON-LAIRD
Psycholinguistics Research Unit and Department of Psychology,
University College London

This experiment was designed to determine whether individuals reason correctly about disjunctive rules. The task consisted in the selection of appropriate instances either to prove, or to disprove, a given disjunctive rule. When the first component of the rule was negated, i.e. when the rule was logically equivalent to implication $(\bar{p} \vee q)$, the selection of appropriate instances was significantly more difficult than when the first component was not negated. The majority of subjects, however, revealed patterns of reasoning which were unstable and labile. The results are discussed in relation to those of previous experiments in which subjects had to reason about conditional sentences.

Introduction

Previous experiments (Wason, 1968) have shown that very few intelligent adults reason adequately about conditional sentences of the form, "if p then q". For example, given a sentence like, "if there is a D on one side of a card, then there is a 3 on the other side", together with four cards bearing respectively D, 3, B, and 7, very few subjects recognized that the 7 was relevant to finding out whether the sentence was true or false. The form of the words, "if p then q," however, is only one way of expressing the concept of implication $(p \supset q)$. It can be expressed in different words using different logical constants, e.g. "there isn't p without q"—$(p\bar{q})$, or in disjunctive form, "either there isn't p or there is q, or both" $(\bar{p} \vee q)$. Is the difficulty of "if p then q" inherent in the concept of implication, or is it due to the words used to express the concept? The present investigation attempts to answer this question by testing the disjunctive form. It should be noted that the appropriate disjunctive form is *inclusive* rather then *exclusive*, i.e. the sentence is true when $\bar{p}$ is true, when q is true, and when both $\bar{p}$ and q are true; it is false only when p and $\bar{q}$ are true. In the present investigation, however, the inclusive inter-pretation was assumed and the words, "or both," were omitted from the material.

Since the propositional calculus counts the following combinations of the conditional as true, pq, $\bar{p}q$, $\bar{p}\bar{q}$, and only the one combination, $p\bar{q}$, as false, it can readily be seen that $(p \supset q)$ and $(\bar{p} \vee q)$ have the same truth table. But although the conditional and the disjunctive form of implication have the same truth table, they have a different meaning. Previous research (Wason, 1968; Johnson-Laird and Tagart, unpublished) has shown that the expression, "if p then q," is construed as true only when both p and q are true. Thus the expression, "either there isn't p or there is q," would appear to be a more accurate description of the truth table.

It is known that the errors made in evaluating negative sentences are considerably greater than those made in evaluating affirmative sentences (Wason, 1959, 1961). Hence in the present investigation the disjunctive form of implication $(\bar{p} \vee q)$ was compared with the unnegated disjunction $(p \vee q)$. $(p \vee q)$ provides a base line against which the difficulty of $(\bar{p} \vee q)$ may be measured. Even though the present task is effectively binary, i.e. $\bar{p}$ takes only one value, it was predicted that more errors would occur under the $(\bar{p} \vee q)$ condition than under the $(p \vee q)$ condition.

In the previous experiments (Wason, 1968), the subjects' task was to select all those values necessary to determine whether a given conditional sentence was true or false. In the present investigation the subject was either told that the sentence was true, or that it was false, and then instructed to select values in order to prove this. Knowledge that the sentence is true or false does not logically affect the correct response—it is the same in both cases. When the sentence is in the form $(\bar{p} \vee q)$ the correct response is to select p and $\bar{q}$; when it is in the form $(p \vee q)$ the correct response is to select $\bar{p}$ and $\bar{q}$. However, when the sentence is known to be true it was considered more likely that the subject would make an erroneous "matching response," i.e. he would select just those values mentioned in the sentence. Similar results have been noted in the attainment of disjunctive concepts by Bruner *et al.* (1956), and in a classification task involving disjunctive sentences by Johnson-Laird and Tagart (unpublished). It was considered, in addition, that this effect would be more pronounced with $(\bar{p} \vee q)$ than with $(p \vee q)$ since it was anticipated that the latter would be too easy to reflect the effect. Hence the following interaction was predicted: the difference between performance in the "true" and "false" conditions would be larger for $(\bar{p} \vee q)$ than for $(p \vee q)$.

The test sentences which were used in this experiment referred to all the stimuli which were presented simultaneously to the subject, i.e. they were in the form, "Every card hasn't a P on one side, or it has a Q on the other side," or "Every card has a P on one side, or it has a Q on the other side." This allowed another factor in reasoning to be examined. There is an asymmetry in proving the rule true and proving it false. Since under the true condition every correct stimulus was consistent with the rule, and under the false condition every correct stimulus was inconsistent with the rule, every correct stimulus has to be inspected to prove the rule true but a single correct stimulus is sufficient to prove the rule false.

Method

Design

There were four basic experimental conditions: $(\bar{p} \vee q)$T, i.e. true, $(\bar{p} \vee q)$F, i.e. false, $(p \vee q)$T and $(p \vee q)$F. The subjects acted as their own controls and were presented with these conditions once, ordered in a different one of each of the 4! permutations.

Subjects

Twenty-four volunteer undergraduate subjects of University College London were individually tested.

Materials

The following four types of lexical material were constructed:

(A) Every card has a number which is Roman on one side, or it has a letter which is capital on the other side.

(B) Every card has a rectangle which is vertical on one side, or it has an angle which is acute on the other side.

(C) Every card has a square which is black on one side, or it has a line which is crooked on the other side.

(D) Every card has a number which is even on one side, or it has a letter which is a vowel on the other side.

Each of these four types of material was, in addition, negated in its first clause so that it corresponded to $(\bar{p} \vee q)$, e.g. "Every card has a number which isn't Roman on one side, or it has a letter which is capital on the other side."

These eight sentences were typed on separate cards (5 × 3 in.). It will be noted that the sentences are not strictly binary, but the stimuli used in conjunction with them made them binary. For each of the sentences eight stimulus cards were constructed consisting of exemplars of p or $\bar{p}$ on one side and q or $\bar{q}$ on the other side. The stimuli used for the four sets of lexical material were as follows:

(A) Roman numbers, Arabic numbers; capital letters, lower case letters.

(B) Vertical rectangles, horizontal rectangles; acute angles, obtuse angles.

(C) Black squares, white squares; crooked lines, straight lines.

(D) Even numbers, odd numbers; vowels, consonants.

Table I shows the logical values on the stimulus cards associated with each of the four experimental conditions. The values given first of all are those on the front of the cards, and those given second in brackets are those on the back of the cards.

TABLE I

The logical values on the stimulus cards and the number of cards associated with each experimental condition

	$p(p)$	$q(p)$	$p(\bar{q})$	$\bar{q}(p)$	$\bar{p}(q)$	$q(\bar{p})$	$\bar{p}(\bar{q})$	$\bar{q}(\bar{p})$	
$(p \vee q)T$	1	1	1	2	2	1	—	—	8
$(p \vee q)F$	1	1	1	—	—	1	2	2	8
$(\bar{p} \vee q)T$	2	1	—	—	1	1	1	2	8
$(\bar{p} \vee q)F$	—	1	2	2	1	1	1	—	8

The explanation for the choice of the particular set of values for each condition is as follows. A necessary constraint is that each of the values p, $\bar{p}$, q, $\bar{q}$, should appear twice on the front of the cards, making a total of eight cards. Thus in the $(p \vee q)T$ condition it is necessary for $\bar{q}(p)$ and $\bar{p}(q)$ to occur twice because inclusion of $\bar{p}(\bar{q})$ or $\bar{q}(\bar{p})$ would render the rule false. A second constraint is that under the false conditions all four correct cards should render the rule false. Thus for the $(p \vee q)F$ condition it is necessary that $\bar{q}(p)$ and $\bar{p}(q)$ be omitted. It will readily be seen that similar principles apply to the stimuli in the $(\bar{p} \vee q)T$ and $(\bar{p} \vee q)F$ conditions. Inspection of the table also shows that the correct stimulus cards are those which are doubly represented in each condition.

Procedure

The first 12 subjects were presented with the lexical material in the order A, B, C, D, and the second 12 subjects in the order D, C, B, A. Under each of the four experimental conditions the appropriate sentence was presented and the eight corresponding stimulus cards placed on the desk in front of the subject in a random array. It was explained to the subject that each sentence was a rule. Each condition had three phases:

(1) The subject was told that the rule was true (false), and that he had to pick out the four cards which he needed to turn over in order to prove this. He was told that there were only four cards which it was necessary to inspect. The subject was given as much time as he needed to think, and then selected four cards.

(2) The experimenter slowly turned over the cards selected by the subject, one at a time, and the subject was asked to tell him as soon as he was sure that the rule had been proved true (false).

(3) The subject was asked why he had not chosen the four remaining cards.

Results

Each subject was given a score from 0 to 4 for each condition according to the number of correct cards which he had selected. The mean scores were as follows: $(p \vee q)T = 2{\cdot}9$, $(p \vee q)F = 3{\cdot}3$, $(\bar{p} \vee q)T = 2{\cdot}6$, $(\bar{p} \vee q)F = 2{\cdot}7$. The prediction that the negative conditions would be more difficult than the affirmative ones was confirmed ($P = 0{\cdot}04$, Wilcoxon test, one-tailed). There was no significant difference between the true and false conditions and the predicted interaction was not confirmed.

Table II shows the frequency of the selection of the logical values for each condition.

TABLE II

Frequency of the selection of logical values for each condition

	$\bar{p}\bar{q}$	pq	$\bar{p}q$	$p\bar{q}$	Others	Totals
$(p \vee q)T$	17	6	0	0	1	24
$(p \vee q)F$	19	3	2	0	0	24
$(\bar{p} \vee q)T$	3	2	6	13	0	24
$(\bar{p} \vee q)F$	4	2	4	12	2	24

The frequencies in Table II, except those classified under "others," refer to the four selected cards, two for each value specified, e.g. the frequency of 17, under $(p \vee q)T$ for $\bar{p}\bar{q}$, means that in 17 cases two values of $\bar{p}$ and two of $\bar{q}$ were selected. The selections classified under "others" consisted of two cases of subjects selecting one value of each of the four values, p, $\bar{p}$, q, $\bar{q}$, and in one case of selecting two values of q and two of $\bar{q}$. The value of χ^2 was computed independently for each condition and was significant in each case. It will be noted that, under all conditions, a correct selection was made more frequently than any other selection. Ignoring truth and falsity, $\bar{p}\bar{q}$ was selected 36 times out of 48 for $(p \vee q)$, and $p\bar{q}$ was selected 25 times out of 48 for $(\bar{p} \vee q)$. Only four out of the 24 subjects were correct for all four conditions.

The number of cards which the subjects needed to turn over in phase 2 before making a decision that the rule was true, or false, was computed independently for the *T* and *F* conditions. Only correctly selected cards were included because under the *F* conditions incorrectly selected cards do not falsify the rule. Each subject had two trials in the T condition and two in the *F* condition, and since a subject may make an incorrect selection on one or both trials, mean values over the two were computed. Table III shows the number of subjects associated with each of these mean values. The correct number of cards to turn over in order to verify the rule is four, and the correct number to falsify it is one. Hence it will be noted that most subjects made a good approximation to the rational strategy.

TABLE III

Mean number of cards turned over to prove that the rule is true or false

Mean number of cards turned over	Number of subjects True	False
4·0	11	3
3·0	1	2
2·5	1	2
2·0	2	0
1·5	0	1
1·0	3	13
Subjects making incorrect selections on both trials	6	3

At the end of the experiment an informal test was made of whether the subjects had been reasoning by inclusive or exclusive disjunction. The remaining cards were turned over by the experimenter to see whether a card which had *both* the values mentioned in the rule was counted by the subject as consistent or inconsistent with the rule. In fact only three subjects gave evidence of reasoning exclusively.

Discussion

It is evident that the element of negation, which is inevitably present in the logical structure of implication, does, to some extent, contribute towards its difficulty. 52·1 per cent of the subjects reason correctly when implication is expressed in disjunctive form, compared with 75 per cent when disjunction is not negated.

It is of interest to observe the types of error made in this experiment. Under the (p v q) conditions matching responses (mirror images of the correct selections) accounted easily for the majority of errors—9 out of 12. On the other hand, under the ($\bar{p}$ v q) conditions such responses accounted for only 10 out of 21 errors. It seems likely that the negative component induces a confusion which manifests itself in a greater variety of erroneous responses. There were 12 matching responses under the *T* conditions as opposed to 7 under the *F* conditions—a trend which is consistent with the predicted interaction between the truth of a rule and its logical form. Contrary to the prediction, however, the difference between the matching responses for the *T* and *F* conditions was more pronounced for (p v q) than for ($\bar{p}$ v q).

When correct selections of cards had been made the reasons given for rejecting the remaining cards tended to reveal an appreciation of disjunction. "Those can be seen to prove at least half the rule." "Those four satisfy the conditions; so I do not need to turn them." When incorrect selections had been made the reasons reflect the erroneous matching principle and, in at least one case, a conflation between disjunction and conjunction. "Those are straight lines and white squares. The rule requires crooked lines and black squares." "Those are already incorrect on one side. I don't know what is on the other side. You just can't do it with four cards." "I took the four which had an even number and a vowel so that if it had the other on the back it would prove the rule true."

A very interesting effect due to the truth of the rule was observed. Four subjects matched on both the true conditions, but three of them made correct selections on both the false conditions. When the rule was true the following were the reasons for rejecting the remaining cards. "You are trying to prove the rule is true and these might disprove it." "Those others are not what is stated in the rule and the rule is true." "I chose the true ones. No matter what is on the other side, these would proved the rule." When the rule was false, these subjects made the correct selections for the wrong reasons. "If I am trying to prove the rule false, then it is obvious you wouldn't choose the ones which prove it correct." "Those others are what is stated in the rule and the rule is supposed to be false." "Those others are half of the rule which I'm trying to disprove."

The results of this experiment, when compared with those of Wason (1968), show that expressing implication as the disjunction, "either not-p or q", makes it easier to grasp than expressing it as the conditional, "if p then q." Just over half the subjects reasoned efficiently about the disjunctive expression, but in the previous experiments only 16·7 per cent of the subjects selected $\bar{q}$. Thus the verbal guise, or description, of a rule has a marked effect on the way it is understood.

The most striking differences, however, are not just quantitative; they are qualitative. When the subjects are given the rule in the form "if p then q," they tend to go wrong in a decisive way. And little that one can do will evidently prevent them falling into error. No such feeling of finality attached to the responses in the present experiment. Sixteen out of the 24 subjects spontaneously doubted that some of their selections were correct. These doubts were expressed on 17 incorrect selections and on 4 correct ones. "I should have chosen the capitals instead of the lower case letters. They would prove the rule false" (selection right, comment wrong). "I was working on the wrong assumption. I could have done it more easily if I had chosen the other four" (selection right, comment wrong). "My mind went blank. Please let me think again" (selection wrong, comment right).

The labile quality of this kind of reasoning seemed to be associated with both the affirmatively and negatively expressed disjunctions. One consequence was that the subjects were ready both to admit error and to gain spontaneous insight into their mistakes after they had made them. Very frequently the act of committal to a choice immediately evoked misgivings. It is as if the "either . . . or" expression itself creates uncertainty. It breaks up the "direction" which seems to be strongly imposed by the conditional, "if . . . then," sentence. With a conditional the individual is likely to be confident but wrong; with a disjunction he is more likely to be unconfident but right. The meaning of a conditional gives no hint of the negation or falsity which underlies its logic. The disjunctive expression makes this element explicit, but this seems to weaken the grounds upon which any inference can be made. In conclusion, a tentative answer may be given to the question posed in the Introduction. There would seem to be inherent difficulties in the concept of implication, but the words in which it is expressed also effect both its ease of understanding and the manner in which it is construed.

References

BRUNER, J. S., GOODNOW, J. J. and AUSTIN, G. A. (1956). *A Study of Thinking*. London.

WASON, P. C. (1959). Processing of positive and negative information. *Q. Jl exp. Psychol.* **11**, 92–107.

WASON, P. C. (1961). Response to affirmative and negative binary statements. *Brit. J. Psychol.* **52**, 133–42.

WASON, P. C. (1968). Reasoning about a rule. *Q. Jl exp. Psychol.* **23**, 273-81.

Received 6 *August* 1968

Q. Jl exp. Psychol. (1969) **21,** 21–28

COGNITIVE CATEGORIES AND MEMORY SPAN

III. EFFECTS OF SIMILARITY ON RECALL

A. F. SANDERS AND J. J. F. SCHROOTS
The Institute of Perception RVO-TNO, Soderberg

The effects of similarity on short-term recall are investigated, by presenting lists of six consonants followed by short lists of consonants and more or less remote cognitive categories, i.e. digits, tones or spatial positions. Some evidence for release from interference, due to dissimilarity of interfering material, is found with consonants and tones, while almost complete release is shown when consonants are followed by spatial positions. Control experiments are included to account for alternative explanations, such as relative difficulty of the various categories, and different memorizing strategies for items, made up of many as opposed to two alternatives. The latter factor appeared to be of particular importance.

Introduction

A very critical issue in the controversy between decay and interference theory to explain forgetting in short-term memory (STM) concerns the effect of similarity on short term recall (Broadbent, 1963; Melton, 1963). It has been repeatedly shown that STM is sensitive to acoustic similarity (Conrad and Hull, 1964) and *not* to semantic similarity, as is usually found in verbal learning (Baddeley, 1966) but this has not been considered as clear evidence for interference theory. Other studies have shown effects of proactive and retroactive inhibition in STM (Sanders, 1961; Keppel and Underwood, 1962; Keppel, 1965; Wickelgren, 1965). The problem here is that a memory span type of experiment is not a pure measure of the contents of the hypothetical STM mechanism, but rather reflects a mixture of short- and more long-term effects. Some parts of a list can resist forgetting much better than other parts (Sanders, 1961) and a list, which is repeated within a test session, gradually shows better recall than non-repeated lists (Hebb, 1961). Therefore, the effects of proactive and retroactive inhibition may be solely due to the long term components.

A strong test between interference and decay theory may be obtained by presenting lists, which consist of successive sublists from more or less remote cognitive categories. An interference theory would not predict substantial forgetting as long as the sublists are both within their respective memory spans and when the categories are remote in the sense that there are almost no mutual associative bonds between elements. A decay theory, on the other hand, would predict the same amount of forgetting, whatever categories are used, since it is assumed that forgetting is mainly dependent on the time-lag between presentation and recall. So the decay theory is supported by data which suggest a constant attention factor (Mackworth, 1964) or a constant STM capacity in the case of recall of multi-dimensional signals (Harrison, 1967).

In an earlier paper (Sanders and Schroots, 1968*b*), an outline was given of a method to determine mutual differences between cognitive categories in an STM setting. It appeared that digits are more related to consonants than tones are, and that spatial positions are most remote from consonants. In the experiments that follow, it will be investigated how far successive sublists of these categories show forgetting in comparison with lists from the same class.

Method and Procedure

The general plan of the experiments was to present lists of six consonants (6C), supplemented with a varying number of items from another category—i.e. digits, tones or spatial positions—or supplemented with more consonants. In the following the supplements will be indicated by RE (remaining elements) when the items are from another category and by RC (remaining consonants) when they are also consonants. The list length of RE and RC ranged from 2 to 6 in the case of digits and tones and from 1 to 5 in the case of spatial positions. Hence, the total list length (6C + RE, or 6C + RC) was minimally 8 and maximally 12 when RE consisted of digits or tones, and ranged from 7 to 11 when RE consisted of visual spatial positions.

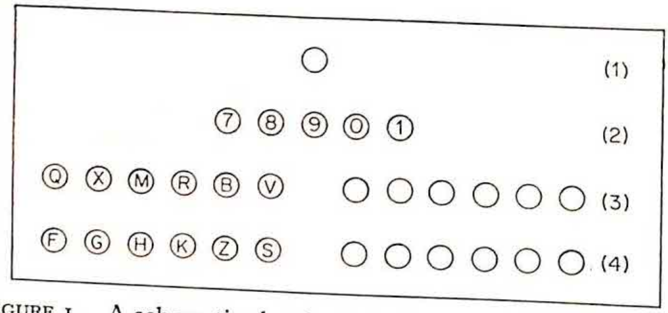


FIGURE 1. A schematic sketch of the visual display (Group III).

For each total list length in a particular condition, ten lists were constructed. All lists were randomly presented at a presentation rate of two signals per second. Ample time was given for recall, which took place after presentation of each list. Subjects were informed each time about length and structure of the next list, i.e. 6C + RE or 6C + RC. The consonants were always from the set (B, F, G, H, K, M, Q, R, S, V, X, Z), which was on continuous display. No consonant was repeated within a list.

Items were only scored as correct when reproduced in the correct serial position. Subjects were informed about this and were instructed to leave open spaces or to guess on locations where items had been forgotten. Three groups of freshmen from the university of Utrecht took part in the experiment. Each group recalled 100 lists, divided in 50 experimental lists (6C + RE) and 50 control lists (6C + RC). Group I consisted of 8 subjects and had digits as RE. The digits were taken from the set 0 through 9, and no digit occurred more than once in a list. Group II contained 7 subjects and had tones as RE. To avoid coding problems, only two tones were used, i.e. a high- and a low-pitched tone (262 and 524 Hz). Recall of both groups was by writing on a response sheet. The tones were symbolized by an upward (↑) and a downward (↓) arrow. Both groups received auditory presentation. All lists were recorded in advance on magnetic tape. All subjects in each group were tested at the same time. Group III consisted of 7 subjects and was given visual spatial positions as RE. A visual presentation technique was needed therefore, the details of which are described elsewhere (Sanders and Schroots 1968*b*; van Doorne and Sanders, 1968). It may suffice to show a schematic sketch of the display, on which the items were successively presented (Fig. 1). The number of items to appear in a list was indicated before presentation by a 1-sec. flash of the appropriate light in row 2 of Figure 1. In case of a 6C + RC list the top row light flashed at the same time, but not in the case of 6C + RE. Consonants were mounted on top of the lights at the left side of the rows 3 and 4, and were only visible during presentation. The lights themselves were mounted on reaction keys. Therefore, a 6C + RE list consisted of a succession of flashes: the first six flashes made consonants appear, the remainder only lights, which had to be retained on the basis of spatial position. The consonants were reproduced by writing them down and the lights by pressing the succession of positions. Each subject was tested individually. The number of correct items was calculated for each subject averaging over the ten lists of the same length and composition. These numbers were transformed to percentage correct recall of the 6C part, and separately, of the RE or RC part. In this way both parts of the list were equally weighted. This exposition outlines only the methods and procedure of the basic experiments. It proved to be necessary to carry out an additional number of control studies, which are described separately at appropriate places.

Consonants and Digits

Results

Mean percentages of correct recall of Group 1 are presented in Table I. An analysis of variance, followed by *post hoc* comparisons of individual means according to the Newman-Keuls procedure showed significant effects of list composition (factor A: $F = 39{\cdot}16$; *d.f.* 1, 19), list length (factor B: $F = 54{\cdot}34$; *d.f.* 4, 91) and subjects (factor D: $F = 24{\cdot}95$; *d.f.* 7, 91). No differences appeared between parts of the lists (factor C), but a significant A × C interaction ($F = 26{\cdot}63$; *d.f.* 1, 91) indicated that performance on either part of the list is related to list composition. The better recall of successive sublists is only reflected in a difference between RE and RC. Finally, a significant C × D interaction ($F = 19{,}36$; *d.f.* 7, 91) showed that performance on 6C vs. RE RC is affected by individual preferences for points of lists.

TABLE I

Mean percentages of correct recall (Group 1, main experiment)

List composition	List length 6 + 2		6 + 3		6 + 4		6 + 5		6 + 6	
Consonants + digits	68·4	88·5	56·5	75·3	61·0	53·7	53·0	38·4	54·3	39·4
Consonants + consonants	65·8	60·5	59·0	52·0	65·0	42·5	48·3	30·3	46·0	27·5

Control experiment and discussion

It would be premature to explain these results as due to categorical differences, in view of the possibility that the effect is a matter of relative difficulty to memorize consonants and digits (e.g. Mackworth 1963; Sanders and Schroots 1968*a*). This alternative requires a control, in which digits are supplemented with consonants. Method and procedure were analogous and ten student subjects took part. The results are in Table II.

TABLE II

Mean percentages of correct recall (control experiment)

List composition	List length									
	6 + 2		6 + 3		6 + 4		6 + 5		6 + 6	
Digits + consonants	79·5	95·5	66·0	90·0	56·0	79·0	51·5	65·0	47·5	65·0
Digits + digits	78·0	91·5	72·0	81·0	58·5	57·5	53·0	48·5	50·5	44·0

Analysis of variance showed significant effects of list composition (factor A), list length (factor B) and subjects (factor D). ($F = 6{\cdot}57$; *d.f.* 1, 9; $F = 11{\cdot}96$; *d.f.* 4·72; $F = 40{\cdot}7$; *d.f.* 9, 12.) Also, a significant A × C interaction was found ($F = 9{\cdot}36$; *d.f.* 1, 9), which was due to better recall of consonants. Therefore, the results are at odds with those of Group 1, which favours the hypothesis of release of interference. The effect is rather slight however and only reflected in better recall of RE. This may not surprise, since consonants and digits are still quite related categories. Furthermore it is still questionable whether the better recall of RE is specifically due to categorical differences. Any structural change—e.g. a short extra time lag—within a list, may provide an anchor, resulting in slightly better recall (Yntema, Worencraft and Klem, 1964).

Nevertheless, the results provide a baseline. When more remote categories are employed as successive sublists more gain should be expected than found here.

Consonants and Tones

Results

Mean percentages of correct recall of Group II are found in Table III. 6C + RE lists are much better recalled than 6C + RC lists (factor A: $F = 234{\cdot}73$; *d.f.* 1, 78)

TABLE III

Mean percentages of correct recall (Group II, *main experiment)*

List composition	List length									
	6 + 2		6 + 3		6 + 4		6 + 5		6 + 6	
Consonants + tones	81·2	93·0	64·8	86·7	59·8	92·5	59·5	80·8	51·8	85·4
Consonants + consonants	62·2	75·0	54·5	59·0	43·0	50·0	45·0	37·4	34·3	34·0

and also list length and subjects differ significantly (factor B: $F=28{\cdot}94$; *d.f.* 4·78; factor D: $F = 25{\cdot}09$; *d.f.* 6, 78). Significant A × B and A × B × C interactions ($F = 3{\cdot}45$; *d.f.* 4, 78; $F = 3{\cdot}29$; *d.f.* 4, 78) indicated that effect of list length is related to list composition and to part of the list. The percentage of correctly recalled tones did not significantly decline as a function of list length, while consonants were better retained in 6C + RE lists (A × C: $F = 38{\cdot}38$; *d.f.* 1, 78).

Discussion and control experiments

Successive sublists have a large positive effect in this condition. Recall of tones remains about equally high as a function of list length, and also consonants are better recalled, when followed by tones; although there is still forgetting at long lists. However, as in the case of consonants and digits, the control is required for effects of relative difficulty. This seems even more urgent since recall of tones may be highly affected by guessing. It was not justified to correct the data of Table III for guessing with the usual formula, requiring randomly distributed errors, since analysis of occurrence of errors in lists of tones showed much more clusters than could be expected on the basis of the probability distribution. However it can be assumed that guessing exerts some effect. A control experiment was carried out therefore, in which lists of tones (8–12), lists of consonants (8–12) and six tones, supplemented with consonants (2–6) were presented. Each condition and list length had six replications making a total of 90 lists. Ten students served as subjects, and the results are presented in Table IV. They show that tones are better recalled than consonants but that best performance is obtained with successive sublists (factor A: $F = 95{\cdot}86$; *d.f.* 2, 18). Hence, better recall of successive sublists cannot be explained by relative difficulty or by guessing, although the data of Table IV show a that part of the improvement can be ascribed to such factors.

TABLE IV

Mean percentages of correct recall (control experiment)

List composition	List length 6 + 2		6 + 3		6 + 4		6 + 5		6 + 6	
Tones + tones	89·6	73·6	84·4	62·8	77·9	52·4	73·6	50·8	67·8	46·8
Tones + consonants	81·7	94·9	80·9	94·8	81·4	90·9	68·9	79·8	71·2	67·6
Consonants + consonants	56·8	62·5	47·6	57·6	52·4	50·3	45·4	47·4	36·3	37·6

An interesting difference between recall of consonants and tones is that better recall of tones primarily occurs in the first part of the list (A × C: $F = 15{\cdot}52$;

TABLE V

Mean percentages of correctly recalled tones as a function of serial position

List length	Serial position 1	2	3	4	5	6	7	8	9	10	11	12
8	93	88	90	90	85	75	68	75				
9	93	90	83	82	82	62	65	57	65			
10	87	82	73	70	75	68	47	45	45	57		
11	68	82	70	75	68	50	52	35	42	58	68	
12	85	67	65	65	63	40	27	30	30	47	55	77

d.f. 9, 72). Analysis of recall as a function of serial position (Table V), shows a consistent high level of recall at the first five or six items, followed by a drop. The recency effect, which is generally found in recall of consonants, only appears at lists of length 11 and 12. Thus, performance in recall of tones, and perhaps in all two-alternatives sequences seems to be better on the first part of the list. It is not improbable that most guessing is found in the second part. At least, the results suggest a difference in memorizing strategy between lists of many and two alternatives.

Control experiment 2

Possible differences in memorizing strategy constitute another alternative explanation of the data of Table III and IV. The test included again three conditions: 6C + RE, 6C + RC and 6C + RE_1, where RE_1 indicates consonants from a set of two alternatives. List lengths (8–12) and conditions had six replications, making a total of 90 lists. Seven students served as subjects. The results are in Table VI. The significant effects were list composition (factor A: $F = 50{\cdot}79$; *d.f.* 2, 12), list length ($F = 37{\cdot}75$; *d.f.* 4, 10), subjects (factor D: $F = 15{\cdot}29$, *d.f.* 6, 104) and A × C ($F = 19{\cdot}23$; *d.f.* 2, 12). *Post hoc* comparisons showed no difference between 6C as a part of 6C + RC and 6C + RE_1, while 6C as a part of 6C + RE was better recalled. RE_1 and RE delivered about equal performance, both being superior to recall of RC.

TABLE VI

Mean percentages of correctly recalled items (control experiment)

List composition	List length									
	6 + 2		6 + 3		6 + 4		6 + 5		6 + 6	
Consonants + tones	76·2	87·0	76·2	76·0	69·0	82·0	67·5	72·8	65·0	70·7
Consonants + consonants (two alternatives)	60·7	97·5	47·7	86·7	45·1	86·2	31·7	76·2	40·0	68·2
Consonants + consonants	53·8	80·0	54·0	61·0	46·5	54·7	35·7	51·4	31·7	43·7

Hence, the results show indeed an effect of changing the number of alternatives within a list: tones are as well recalled as consonants, consisting of 2 alternatives. On the other hand, 6C lists are only better recalled if they are followed by tones. This seems to be the only genuine effect of the categorical differences. It shows that, using consonants and tones, only partial release from interference is found.

Consonants and Spatial Positions

The results of Group III are presented in Table VII. The statistical analysis showed significant effects at factors A, B and D ($F = 27{\cdot}37$; *d.f.* 1, 6; $F = 21{\cdot}76$; *d.f.* 4, 76 and $F = 11{\cdot}23$; *d.f.* 6, 76) and at A × B ($F = 3{\cdot}26$; *d.f.* 4, 76) and A × C ($F = 7{\cdot}69$; *d.f.* 1, 76). The effect of list length was much stronger at 6C + RC. 6C as a part of RE did not significantly decline, when RE increased from one to five items. RE declined only at list length 5. So, the results of Group III show that 6C + RE lists are much better recalled. It must be especially noted that both 6C and RE show no sizeable decrement, when list length increases. The decline of RE at list length 5 is in line with data on memory span for spatial positions (Sanders, 1968).

TABLE VII

Mean percentages of correctly recalled items (Group III, *main experiment)*

	List length									
List composition	6 + 1		6 + 2		6 + 3		6 + 4		6 + 5	
Consonants + positions	85·7	97·0	85·5	95·5	82·7	85·7	81·7	84·7	78·3	72·7
Consonants + consonants	85·7	76·0	74·8	65·6	67·7	45·3	62·7	42·7	60·1	34·6

It does not appear necessary to carry out control studies as was done previously. First, spatial positions have a lower span than consonants (Sanders, 1968), so that the argument of relative difficulty is not valid. Second, the number of alternative consonants and positions were equal. It seems justified, therefore, to ascribe the almost complete release from interference to the very marked categorical differences. When 6C + RE had a length of 10, 42·9 per cent of the lists were correctly recalled, making an average span of about 10 items, instead of the magical number seven (Miller, 1956). It may be that this large effect is related to different coding—i.e. acoustic with consonants and visual with spatial positions—reducing the mutual interference. Still, there is no complete additivity of memory span, since both RE and 6C may be slightly below the level that is found when they are measured alone. This may constitute a minor decay factor.

General Discussion

The results, taken together, show considerable evidence for improved immediate recall, when lists are made up of successive sublists from most remote categories. The effect is small and questionable at consonants and digits, more clear with consonants and tones, and approaches additivity of memory spans when consonants are followed by spatial positions. This is not in line with Mackworth's theory of immediate recall, as determined by a constant attention factor (Mackworth 1964) or with Harrison's suggestion of a constant STM capacity when multidimensional signals are to be recalled (Harrison, 1967). The results with multidimensional signals may be due to order of recall. It is probable that when recall is done per item—i.e. all attributes of one item followed by those of the next one—much less correct recall is obtained than when subjects recall categorically—i.e. one attribute of all items followed by the next attribute (Sanders and Schroots, 1968*b*). The latter strategy might enable subjects to recall all 16 values, when four items with four attributes each are presented, as was found on occasions by Harrison.

The data are consistent with a conception of STM in terms of short term activation of pre-existing units and associative bonds in a dictionary store, where materials from more remote classes are less mutually interfering. A decay theory might still be correct, if the assumption of single channel transmission of items does not hold when remote categories are involved. If it is possible to rehearse during memorizing of items of the other class, a decay theory would not predict much forgetting. This possibility suggests some interesting experiments, both on STM and more typical topics from single channel theory such as the psychological refractory period.

References

BADDELEY (1966). Short term memory for word sequences as a function of acoustic, semantic and formal similarity. *Q. Jl exp. Psychol.* **18**, 362–5.

BROADBENT, D. E. (1963). Flow of information within the organism. *J. verb. Learn. verb. Behav.* **2**, 34–9.

CONRAD, R. and HULL, A. J. (1964). Information, acoustic confusion and memory span. *Brit. J. Psychol.* **55**, 429–32.

DOORNE, H. VAN and SANDERS, A. F. (1968). A programmable signal and response processor. *J. Behav. Instruments (PSARP* **1**, 29–32).

HARRISON, G. J. (1967). Some additive results in short term memory. In *Attention and Performance*, A. F. SANDERS (Ed.), *Acta Psychol.* **27**, 306–15.

HEBB, D. O. (1961). Distinctive features of learning in the higher animal. In *Brain mechanisms and Learning.* Pp. 37–46, J. F. DELAFRESNAYE (Ed), England: Oxford University Press.

KEPPEL, G. (1965). Problems of method in the study of short term memory, *Psychol. Bull.* **63**, 1–13.

KEPPEL, G. and UNDERWOOD, B. (1962). Proactive inhibition in short term retention of single items. *J. verb. Learn. verb. Behav.* **1**, 153–61.

MACKWORTH, J. F. (1963). The relation between the visual image and post perceptual immediate memory. *J. verb. Learn. Behav.* **2**, 75–85.

MACKWORTH, J. F. (1964). Interference and decay in very short term memory. *J. verb Learn. verb. Beh.* **3**, 300–08.

MELTON, A. W. (1963). Implications of short term memory for a general theory of memory. *J. verb. Learn. verb. Behav.* **2**, 1–12.

MILLER, G. A. (1956). The magical number seven, plus or minus two: Some limits on our capacity for processing information, *Psychol. Rev.* **63**, 81–97.

SANDERS, A. F. (1961). Rehearsal and recall in immediate memory. *Ergonomics* **4**, 25–34.

SANDERS, A. F. (1968). Short term memory for spatial positions, *Ned. Tijdschr. Psychol.* **23**, 1–15.

SANDERS, A. F. and SCHROOTS, J. J. F. (1968*a*). Cognitive categories and memory span: I. Shifting between categories. *Q. Jl exp. Psychol.* (**20**, 370–72).

SANDERS, A. F. and SCHROOTS, J. J. F. (1968*b*). Cognitive categories and memory span: II. The Effect of temporal vs. categorical recall. *Q. Jl exp. Psychol.* (**20**, 373–9).

WICKELGREN, W. A. (1965). Short term memory for repeated and non-repeated items, *Q. Jl exp. Psychol.* **17**, 14–25.

YNTEMA, D. B., WORENCRAFT, F. T. and KLEM, L. (1964). Serial recall of digits presented at very high rates. Paper presented at meeting of Psychonomic Society, Niagara Falls 1964.

Received 27 June 1968

Q. Jl exp. Psychol. (1969) **21**, 29–38

HYPOTHESIS FORMATION AND PROBLEM COMPLEXITY

IVANA MARKOVÁ

Psychological Laboratory, Cambridge

The present study examines the relationship between hypothesis formation and the degree of complexity of the problem which has to be solved. Ninety subjects played a simple arithmetical game with the investigator. They were divided into two groups according to the type of problem given. The hypotheses formed by the subjects in the course of the game were classified under two headings; namely, H_{I}, trying to form a general and comprehensive strategy on the basis of an overall evaluation of the problem; and H_{II} based on a more detailed logical analysis and leading to a strategy with sub-goals. Hypotheses of the first type prevailed in the simpler versions of the problem, those of the second type were more frequent in the more complicated versions. If the number of alternative strategies is small, subjects are not forced to develop an abstract formulation of the strategies, and continue to operate with concrete strategies (e.g. I always take two chips). When there are many alternative strategies, subjects are forced to use more abstract analyses of the problem which appears to involve formulation of sub-goals. The consequences of both types of hypotheses for problem solving are discussed.

Introduction

Behaviour during problem solving appears to be controlled by plans (Miller, Galanter and Pribram, 1960) or hypotheses (Wason, 1960) or strategies (Bruner, Goodnow and Austin, 1956). In the process of problem solving a subject may select one of many possible procedures for analysing and processing information, which he has either obtained beforehand in the form of instructions or which he has acquired in the process of learning. Scientific inference from hypotheses and finding criteria of plausibility of hypotheses is of concern in current philosophy of science (Barker, 1957). Making inferences from hypotheses in problem solving, and its relationship to scientific thinking, has been discussed by Wason (1960) who finds that "even" intelligent adults do not readily adopt a scientific attitude to a novel problem" (Wason, 1967). These findings raise the question of whether a "non-scientific attitude" has any connection with the structure of the problem given. Broadly speaking, the present investigation focusses on the level of generality of hypotheses formed by the S. and, in particular, on the relationship between the hypotheses and the complexity of the problem.

Variants of the game of "NIM" game were used in the investigation.[†] "NIM" is a game for two and, according to game theory classification, it involves a zero sum and complete information. It is played as follows. We have a pile of chips from which both players take alternately either one chip or several chips according to the rules fixed in advance; the player left with the last chip is the loser. This game, in another variation, was investigated by Peterson (1920) and Waters (1928).

[†] A description of this game was given in a paper read at XVIII International Congress of Psychology in Moscow 1966 (Marková, 1966).

The number of chips at the start is determined by the formula: $nx + 1$, where $n =$ positive integer, $x = k + 1$, and $k =$ maximum number of chips which may be taken. If both players know the optimal strategy, then the winner is always the player who plays second. Here optimal strategy is introduced only intuitively: each player is allowed to take at least one chip and not more than $x - 1 = k$ chips in each move; thus the second player can always complete the number of chips taken by the first player up to the number $x = k + 1$. Suppose, for example, that the total number of chips is 13. The rule specifies that, in each move, the player is allowed to take 1, 2, or 3 chips. The second player, if he plays the optimal strategy, must always complete the number of chips taken to a multiple of four.

Definitions

(1) *Hypothesis.* In our work we start from the assumption that, in theoretical problems, the subject formulates hypotheses as propositions (e.g. hypothesis: On Mars there is life), whereas in practical problem solving (as in the NIM game, where the subject is supposed to perform an action to reach his goal), the subject *transforms hypotheses into imperatives* in his verbalization and in his thinking. On the basis of this assumption, we may infer hypotheses from the statements expressed by subjects, e.g., the hypothesis "if I take an odd number of chips in each draw I shall win" is transformed by subject directly into a prescription for action, which is verbalized as "take an odd number of chips in each draw."

(2) *Set.* We shall call each ordered k-tuple of ordered pairs of digits, which might make up an actual game, a set. Thus,

$$((i_1, j_1) \ldots\ldots\ldots\ldots (i_k, j_k)).$$

$i_u =$ number of chips taken by the investigator in the uth round (one round is each ordered pair of draws of investigator and the subject, who is always the second player). And,

$j_u =$ number of chips taken by the subject in the uth round.

The digit k is determined in the following way:

k = the only x such that

$$\left(\sum_{m=1}^{x} i_m + \sum_{n=1}^{x} j_n = y \right).$$

(3) *Strategy.* Strategy is such prescription for action as enables some potential sets to be rejected and other sets realized. Each strategy therefore "cuts out" a certain definite sub-class of sets.

(4) *Problem.* The problem given to subject is to find a rule which determines the value of j_u in dependence on i_u with the result that one of the following sets arises:

$$\begin{array}{l} ((i_{1k}, j_{1k}) \ldots\ldots\ldots\ldots (1, 0)) \\ ((i_{2k}, j_{2k}) \ldots\ldots\ldots\ldots (1, 0)) \\ \quad\vdots \qquad\qquad\qquad\quad \vdots \\ ((i_{nk}, j_{nk}) \ldots\ldots\ldots\ldots (1, 0)). \end{array}$$

This is a class of sets where i_u is arbitrary and j_u is determined by the rule prohibiting the investigator from making an optimum draw.

(5) *An optimum draw* by the investigator in the uth round of a game would be the draw by which he completes the number of chips so far taken to a multiple of n.

(6) *The optimum game* for the subject would be the game in which the subject completes the number of chips taken to a multiple of n in each round.

(7) *The optimum strategy* of the subject would be such a strategy which eliminates those games which are not the optimum games of the subject.

(8) *The solution* of the problem given to the subject would be finding the rules of the optimum strategy.

Strategies are formulated in the statements of subject's as prescriptions for action. A hypotheses or a sequence of hypotheses may be ascribed to each strategy. Each strategy may be connected with several different hypotheses. For example, let us have the strategy exemplified by use of the set: $\langle\langle 3,1\rangle \langle 1,3\rangle \langle 1,3\rangle \langle 1,0\rangle\rangle$. From this strategy, we cannot reliably infer which hypothesis the subject played—whether according to: to take only odd chips; the first draw one, others draws three; to complete the number of chips taken in one round to an even number; to complete the number of chips taken in one round to four; etc.

Method

Procedure

There were two series of experiments. Our assumption was that the complexity of the game depended on the total number of chips used in the game and also on the number of chips which each player could take in each draw. In each series there were five variants of the NIM game. Four of them were actually played by manual manipulation of chips, and one variant was considered only theoretically as a test of the strategy developed.

First series. The games were in the following sequence: G^{13}_{3}, G^{13}_{4}, G^{16}_{5}, G^{19}_{6} (actual play); G^{50}_{7} (theoretical variant).

G—stands for the game. The subscript denotes the total number of chips for both subjects which must be taken in each round if the second player plays perfectly (optimum strategy). The superscript denotes the total number of chips in the game.

Second series. The games were in the following sequence: G^{19}_{6}, G^{16}_{5}, G^{13}_{4}, G^{13}_{3} (actual play); G^{50}_{7} (theoretical variant).

The following rules were imposed. If after 20 sets or 60 min. the subject was unable to arrive at a hypothesis which determines the class of optimum sets, then the game was terminated. The investigator always started off the game, playing as well as the situation allowed. For each subject the number of chips taken by the investigator at the beginning of each set was as follows.

G_3^{13}: 1 2 1 2 1, etc.
G_4^{13}: 1 3 2 1 3, etc.
G_5^{16}: 1 3 2 4 1 3 2 4, etc.
G_6^{19}: 1 3 5 2 4 1 3 5, etc.

This arrangement put all the subjects on equal footing, and reduced the number of sets.

The subjects were 90 children, 15 years old at Prague comprehensive schools. They were selected at random among pupils with the highest marks in maths. The subjects were tested individually in a room in the school. They were divided into two groups of 45 subjects equalized according to score on the Raven matrices. The scores gave an I.Q. range of -115-132. The first group had 23 boys and 22 girls, the second 22 boys and 23 girls.

Instructions for the first game were as follows. "Here is a pile of chips. The number of chips is 13 (first series)—19 (second series). We shall both be taking chips from this pile, one after another. I shall be the first. Both you and I can take 1 or 2 chips (first series)—1 to 5 chips (second series) from the pile. The one left with the last chip from the pile loses the game. Do you understand? The game will be repeated 20 times. I shall always start. While you play, I would like you to think aloud. Please say all that comes into your mind as you try to solve the problem. Try to say everything even if you think that it has nothing to do with the game. Do you understand my instructions?" Where the subject did not understand, the instructions were repeated in the same words.

Instructions for the second, third and fourth game were as follows:

"Now here is a pile of 13 (first series)—16 (second series) and we are allowed to take 1, 2 or 3 chips (first series)—1 to 4 chips (second series) from the pile. As before, we shall both be taking chips from this pile one after another and I shall always start first."

Instructions for the fifth (theoretical) game were as follows.

"Now imagine that there is a pile of 50 chips. In each draw you are allowed to take 1 to 6 chips. How would you manage to win each game if I again always start first?"

This theoretical variant was given at the end of each series only to the subjects who successfully solved the actual games.

Scoring criteria

We can distinguish logically two types of hypotheses:

Hypotheses of the first type, H

Definition	*Typical verbalization*
When (and only when) I carry out such and such sequence of moves, I shall win.	I shall carry out the following sequence of moves. . . .

It is obvious that in hypotheses of this type the subject tries to determine simultaneously all the sets whose final round is $<1, 0.>$ As the H_I governs the whole action of the subject, it is necessary, if the hypothesis proves false, to substitute it with another, completely new hypothesis. Hypotheses H_I are based neither on the analysis of problem situation nor on the analysis of the goal.

Hypotheses of the second type, H_{II}. This is really a sequence of hypotheses. This sequence consists of individual hypotheses that determine individual sub-goals only. The last hypothesis is in terms of an actual draw.

Definition	*Typical verbalization*
(1) When (and only when) I reach the situation A_1 I shall win.	(1) I shall reach the situation A_1.
⋮	⋮
(2) When (and only when) I reach the situation A_m I shall reach the situation A_1.	(2) I shall reach the situation A_m.

The hypotheses formed by the subjects in the course of the games were classified into the two types distinguished above, i.e. hypotheses of the first type, H_{I}, trying to form a general and comprehensive strategy on the basis of an overall evaluation of the problem; and hypotheses of the second type, H_{II}, based on more detailed logical analysis and leading to a strategy with sub-goal aims.

Classification into these two categories was made on the basis of the transcription of actual utterances by subjects in the course of the experiment. Individual moves both of the subject and the investigator were also recorded, but classification of hypotheses was made only on the basis of the verbalizations. This appears to be valid (Markova, 1964). The first two or three sets of the first game were mostly devoid of any verbailization: statements of the subjects were here of the type "I do not know why I have made this move" or " I will try this or that for the time being": the strategy is chaotic and nothing can be extracted from it that would correspond with H_{I} or H_{II}. The stage of problem solving beginning, on average, with the third set of the first game is characterized by statements which can be classified as hypothese H_{I} or H_{II}.

If more than half of subjects verbalizations in the series of four actual games were of character H_{II}, his record sheet was classified as an overall example of H_{II}, and vice versa for H_{I}. These proportions were taken from, on average, five statement per first game and three statements per second, third and fourth game; but it was also true that subjects who used H_{I} produced more verbalizations than subjects who used H_{II}. Changeovers in hypothesis type could be observed in the data. For deciding that such changeovers had occurred, the same criterion, one-half, was applied to each of the four games. The changeovers are shown in Tables I and II.

TABLE I

First group

Changeovers	G_3	$G_3 \rightarrow G_4$	G_4	$G_4 \rightarrow G_5$	G_5	$G_5 \rightarrow G_6$	G_6
$H_{\text{I}} \rightarrow H_{\text{II}}$	0	0	1	2	2	1	0
$H_{\text{II}} \rightarrow H_{\text{I}}$	0	0	0	0	0	0	0
Σ	0	0	1	2	2	1	0

TABLE II

Second group

Changeovers	G_6	$G_6 \rightarrow G_5$	G_5	$G_5 \rightarrow G_4$	G_4	$G_4 \rightarrow G_3$	G_3
$H_{\text{I}} \rightarrow H_{\text{II}}$	4	1	0	0	0	0	0
$H_{\text{II}} \rightarrow H_{\text{I}}$	0	0	0	0	0	0	2
Σ	4	1	0	0	0	0	2

While not significant these results show that changes tended to be in the direction $H_{\text{I}} \rightarrow H_{\text{II}}$, and to be related to the complexity of the game rather than to the stage of the experiment.

Quantitative Results

Two questions are important. First, are the relative probabilities of hypotheses of types H_{I} and H_{II} different for the two experimental groups? Second, is there any connection between the type of hypothesis selected and the successful solution of the game?

The data relevant to the first question are contained in Table II, which shows the number of subjects in each group choosing each type of hypothesis. The relative probabilities of H_I and H_{II} may be said to be different in the groups if Table III is significantly non-homogeneous. The likelihood ratio test for homogenity in contingency tables is discussed by Wilks (1935). The test statistic is

$$\hat{I}(Y_1 : Y_2) = \log \frac{P(\text{data}/Y_1)}{P(\text{data}/Y_2)}$$

where Y_2 is the hypothesis of homogeneity, (i.e. $P(H_I/\text{group } 1) = P(H_I/\text{group } 2) = \alpha$, say) and Y_1 is the alternative hypothesis which permits $P(H_I)$ to assume a different value for the group. Maximum likelihood estimates are used for the parameters implicit in Y_1 and Y_2 and the test statistic then becomes

$$2\hat{I}(Y_1 : Y_2) = \sum_{ij} x_{ij} \log x_{ij} - \sum_{i} x_{i\cdot} \log x_{i\cdot} - \sum_{j} x_{\cdot j} \log x_{\cdot j} + N \log N \qquad (1)$$

where *i.i.* (= 1,2) refer to rows and columns in Table III, respectively, x_i is a row total and $x_{\cdot j}$ is a column total and N is a total number of observations in the table.

$\hat{I}(Y_1 : Y_2)$ is also the estimate of the contingent uncertainty or information transmitted in Table III (see Garner, 1962). As N becomes large, $2\ \hat{I}(Y_1 : Y_2)$ is distributed asymptotically as χ^2 with 1 *d.f.*

It happens that $2\ \hat{I} = 28{\cdot}74$ and this is significant at the 0·005 level.

TABLE III

Frequency of H_I and H_{II}

Experimental group	Type of hypotheses H_I	H_{II}	Σ
First (complexity increase)	37	8	45
Second (complexity decrease)	12	33	45
Σ	49	41	90

The same statistical method was used to answer questions about the relationship of hypothesis type to success in solving. The relevant data are shown in Tables IV and V for the two groups. Each subject's score may appear in any one of four cells

TABLE IV

Frequency of H_I and H_{II} in the first group

Type of hypotheses	Unsolved G_3—G_6	Solved G_3—G_6	Σ
H_I	19	18	37
H_{II}	3	5	8
Σ	22	23	45

according to the type of hypothesis selected and whether the problem is solved. If the selection of H_{I} and the solving of the problem are independent events, then the table will be homogeneous (hypothesis Y_2); otherwise, the data are a sample from some unspecified multinominal distribution (hypothesis Y_1). Equation (1) gives $2\hat{I}(Y_1 : Y_2) = 0{\cdot}51$ with 1 *d.f.* for Table IV, which is not significant, and 3·87 with 1 *d.f.* for Table V, which is just significant on 0·05 level.

TABLE V

Frequency of H_{I} and H_{II} in the second group

Type of hypotheses	Unsolved G_3—G_6	Solved G_3—G_6	
H_{I}	9	3	12
H_{II}	14	19	33
Σ	23	22	45

While Tables IV and V involves data from actual games, Table VI involves data from the theoretical game. The theoretical game, G_7, was played only by those 45 subjects who were successful in actual games. The same statistical method was used.

TABLE VI

Frequency of H_{I} and H_{II} hypotheses in G_7 game

Type of hypotheses	Unsolved G_7	Solved G_7	Σ
H_{I}	11	10	21
H_{II}	0	24	24
Σ	11	34	45

Equation (1) gives $2\hat{I}\ Y_1 : Y_2 = 20{\cdot}99$ with 1 *d.f.* for Table VI, which is significant at the 0·005 level.

Summarizing the statistical results we can see that the relative probabilities of hypotheses of types H_{I} and H_{II} are different for the two experimental groups: in the first group (complexity increasing) the H_{I} type prevails while in the second group (complexity decreasing) H_{II} is the predominating type. For the actual games there is no statistically significant relationship between the type of hypothesis formed and the probability of succesful solution for the first group, while in the second group the dependence is just significant. However, the connection between the type of hypothesis employed and success is highly significant for the theoretical game: only H_{II} leads to the level of generalization required for more complex games.

Qualitative analysis of protocols

Hypotheses of the first type contain an element of chance associating them with trial and error methods, because they are not based on analysis of the problem situation. This points to their "mechanical" character. If the situation is obvious in all actual variants, H_I is viable. Yet subjects may play an optimal game without understanding the structure of the problem; in each new variant of the game he must look for a new hypothesis because the hypothesis verified in one variant turns false in another one.

If we come across a sequence of examples of H_I, we shall find that the relationship between individual hypotheses is very loose and purely external: it may be dependent either on time or an associations according to rules of the contrast, analogy, arithmetic series, etc.

The subjects formed the hypotheses H_I according to their failure in a set. If they were unsuccessful and lost the set, they tried to arrive at another hypothesis not on the basis of analyzing the situation but intuitively only. This explains why, in general evaluation, no criteria were found by which hypotheses of the H_I type could be arranged into a system.†

On the other hand in hypotheses of the second type the individual hypotheses only aim at the sub-goals and, appropriately, transform the original problem into another one. A sequence of hypotheses determines the whole action. The first sub-goal is reached quite empirically by the analysis of the last-but-one round. The situation here is viable as only small numbers of chips to be taken are left. Having solved this situation, the subject tries to find further sub-goals by recursion. When, for example, the subject has to solve the problem G_4 with 13 chips, he transforms it in the problem of how to solve the situation with nine chips, because he finds that if he arrives at the situation with four chips to be taken, he will always loose. But the problem with nine chips is transformed in to the sub-problem with five chips and, at last, he finds that the first move is decisive and that he must complete the chips taken to a multiple of four.

Within hypotheses of the second type some subjects are not able to arrive at the right sub-goal, or they cannot complete the chain between individual sub-goals and an actual draw. Or they might oscillate between two sub-goals, one of which was false. Example: Game G_6: "I must reach the number 13" (wrong). "I must not reach the number 13" (right) "I must reach the number 13," etc. Sub-goals are of two kinds: to be reached by subjects (positive) or to be avoided by the subject (negative). Occasionally sub-goals were determined too narrowly or too broadly, i.e. too few numbers to be reached or to be avoided were determined and vice versa for too broad. In the example of G_3: "I must reach 6 to win." This is too narrow. Six or five would be connect. "I must not reach seven or eight to win," is too broad, because eight does not matter.

---

† *Examples of protocol*

Hypothesis 1: Game G_6; female M.M. Group II, Raven Matrices I.Q.: 121.

(a) I will try to take one chip more than you.
(b) Now I take a half less than you.
(c) Now I will take one chip less than you.
(d) Now I will always take one chip.
(e) I will take the same number as you.
(f) Now whenever you take the highest number of chips I shall take the smallest.

Hypothesis 2: Game G_6 was played: male R.H., exp. group II, Raven 127.

(a) I see, when seven are left I will lose. I must try to have my turn when there are six left.
(b) It is the number seven that is important. How to get to it? I must play in such a way so that the number is for you.
(c) Thirteen chips remain, it depends how many you take. Whatever you take the seven will remain to you.
(d) What will be the situation if I take four at the beginning? You can take one to five No. I will not take four, I will take one for you to get 13.
(e) I do not know yet if I can win every time. I understand it already but I must not make a mistake. Whatever you take at the beginning I must complete the number to such an amount that 13 will remain for you and then seven.

To play well a subject has to realise, at each draw, both which number of chips has to be avoided and which number should be achieved. Some of the less able subjects confused these two numbers. The successful subjects usually determinated the negative subgoal too broadly (e.g. to determine several numbers as decisive for winning although this did not matter).

Sub-goal solving of the problem by means of H_{II} hypotheses brings the subject to the understanding of the structure of the game. This is proved by the fact that almost all subjects who solved G_6 were also successful in G_7 (there were frequent comments such as: "now I could solve any game of that type").

Discussion

I shall try to explain why, in the first group, the H_I hypotheses prevailed while, in the second group, the H_{II} were used: the groups were equalized on the bases of results in maths and in Raven matrices. Success or failure in solving the actual games cannot be related to the type of hypotheses applied. In our opinion they depend rather on the stimulus situation and on structural relationships inherent in the problem. In such a simple game as G_3 a less ambitious approach is possible: in view of its simple rules, it is advantageous for the subject to try all the possible combinations of the set ("to take the same as you", "to take the opposite to you" and similar combinations) rather than to attempt the goal analysis. If the subject wanted to exploit the same possibilities while playing G_6, he would have to formulate a large number of H_I hypotheses which is obviously less effective than formulating the H_{II} type. As soon as the subject accepts this assumption and begins to form hypotheses of a certain type, his strategy becomes stable and he adheres to formation of a certain type of hypotheses even if it is no longer advantageous, i.e. he tries to solve the variant G_3 by means of sub-goal analysis because earlier he had sucessfully applied this analysis to G_6 strategy; on the other hand he solves more ambitious variants by means of H_I, if he made use of it for G_3 before.

Newell, Shaw and Simon (1959) have suggested a means-end analysis superficially similar to our sub-goal analysis based on H_{II} hypotheses. There is a certain analogy between their problem situation and that in our experiment. The problem of regular transformation of any expression A into any expression B, deducible form A, corresponds to our problem of how to transform (by playing) the initial position of G_n into the "winning position," in whatever way the investigator is going to move, and both are thus sub-goal analyses. In both cases there is a "diminishing of a difference": in our case the mth move narrows and finally nullifies the difference between the sub-problems before and after and after the mth move.

However, there are differences as well. The necessity of recursion to determine the whole action in case of H_{II} is not analogous with Newell. The main difference is that the Newell general problem does not enable us (with the given list of operators unlike the Hao-Wang system) to find the algorithm which would solve the problem. Our general problem on the other hand, is strictly solvable: this was revealed in its different presentations: the subject in our experiment succeeded in solving the individual assignment only after having solved the whole general problem, i.e. not only after having won an individual set. In the Newell experiments, the subject succeded in solving the assignment when one of the instances had been solved, i.e. the given expression was transformed into another given expression. This means that the subject was not requested to reach the general principle of regular transformation: this general "principle" was installed,with a certain limitation, into the machine simulating the behaviour of the Newell problem solver and it did not function, and even could not function, in all cases. These differences restrict the analogy between the results from the Newell experiments and the results of our experiment, and our results may not be found in other classes of problem.

I am greatly indebted to Dr. M. Haggard for his help with the manuscript and in clarifying ideas and to Dr. D. Laming for statistical advice.

References

BARKER, F. (1957). *Induction and Hypothesis*. New York: Ithaca.

BRUNER, J. S., GOODNOW, J. J. and AUSTIN, G. A. (1956). *A Study of Thinking*. London.

GARNER, W. R. (1962). *Uncertainty and Structure as Psychological Concepts*. New York: Wiley.

MARKOVÁ, I. (1964). *Acquisition of optimum strategy in a game as a function of the instructions*, unpublished M.S. thesis, University of Prague.

MARKOVÁ, I. (1966). A classification of heuristic procedures in games. *Proc. XVIIIth Int. Congr. of Psychol.*, Moscow.

MILLER, G. A., GALANTER, E. and PRIBRAM, K. H. (1960). *Plans and the Structure of Behaviour*. New York: Wiley.

NEWELL, A., SHAW, J. S. and SIMON, H. A. (1959). Report of a general problem solving program. *Proc. of Information Processing, UNESCO*, pp. 256–65.

PETERSON, J. C. (1920). The higher mental process in learning. *Psychol. Monogr.* **28**, No. 7 Whole No 129.

WASON, P. C. (1960). On the failure to eliminate hypotheses in a conceptual task, *Q. Jl exp. Psychol.* **12**, 129–40.

WASON, P. C. (1967). Reasoning. In *New Horizons in Psychology*. London: Penguin Books.

WATERS, R. H. (1928). Influence of tuition upon ideational learning. *J. Gen. Psychol.* **1**, 534–49.

WILKS, S. S. (1935). The likelihood ratio test of independence in contingency tables. *Ann. math. Statist.* **6**, 190–6.

Received 5 July 1968

Q. Jl exp. Psychol. (1969) **21**, 39–42

RECALL OF DIGITS PROJECTED TO TEMPORAL AND NASAL HEMI-RETINAS

H. SAMPSON

Department of Psychology, University of Auckland

Fourteen right-handed, right eye dominant subjects recalled digits when different ones were projected simultaneously to either temporal or nasal retinas. The principal findings were: (a) Recall of digits projected to nasal retinas was significantly better than when projected to temporal retinas; (b) information projected to the right eye was recalled significantly better than that projected to the left eye. It is shown that the relative ineffectiveness of the recall of input from the *non-dominant eye* can be attributed almost wholly to the relative inefficient recall of digits projected to the *left temporal* retina.

Introduction

Sampson and Horrocks (1967) have shown that if there is competition between inputs from each eye when the subject is required to recall digits presented visually, it is resolved in favour of the dominant eye. In their experiments, different digits were projected simultaneously to each eye, and subjects recalled those projected to the dominant eye more accurately than those projected to the non-dominant eye.

A problem arising from this observation of theoretical interest, is to explain the relative ineffectiveness of the *non-dominant eye*. In view of the well-known partitioning of the retina into nasal and temporal regions, each with different central projections, it is possible that the result obtained because information projected to either one of these regions or perhaps both, is recalled less effectively than when it is projected to corresponding regions of the dominant eye. To examine this possibility, different digits were projected simultaenously to either temporal or nasal hemi-retinas in the present experiment, and the consequences of the resulting binocular rivalry for short-term memory assessed.

Method

Subjects

These were 14 male university students. Their ages ranged from 19 to 28 years (mean = 22 years) and they were all right-handed and right eye dominant. Handedness was determined by an individually administered questionnaire composed of 20 items suggested by Humphrey (1951) as sensitive indices of manual preference, and three additional items designed by Gillies, MacSweeney and Zangwill (1960). Sighting preference was determined by the finger aiming test (Buxton and Crosland, 1937) and use of a card with a 4 cm. square hole in it. An AOA chart was used to check visual acuity, and each subject had normal and equal acuity in each eye.

Apparatus

A magazine-load automatic projector was connected to two interval timers permitting control of stimulus duration and interval between stimuli. Slides were "gravity-fed" into the projector from a vertically placed slide holder. The timers controlled the shutter mechanisms of the projector. A day-view screen upon which stimuli for each eye were projected was placed 1 m. from the subject's eyes. The visual stimuli were arabic digits from 1 to 9, and they were projected in pairs either side of a fixation point so that in terms of visual angle, the centre of each digit was 5° from the fixation point, and a circle of 6° took in each pair of digits entirely. All stimuli were mounted in pairs on 5·1 cm. × 5·1 cm. slides. The fixation point was permanently mounted on the day-view screen. The subject's head was held in a head-rest, and he wore a trial frame when viewing the stimulus pairs. The trial frame's eye pieces were fitted with half occluders which meant that after the subject was seated and had fixated the dot on the screen the occulder could be rotated to either nasal or temporal position and adjusted so that the subject could see the fixation point and only one digit with each eye. Two viewing conditions were used:

Condition I: the occluder for each eye in the temporal position so that the stimulus is projected onto the *temporal* retina.

Condition II: the occluder for each eye in the nasal position so that the stimulus is projected onto the *nasal* retina.

Procedure

The 14 subjects were divided into two groups. Odd numbered subjects viewed stimuli temporally and then nasally. Even numbered subjects viewed stimuli nasally, then temporally. Each subject was given 20 trials under each viewing condition. A trial consisted of four pairs of digits followed by a red flash on the screen signalling the end of a trial. The combination of digit pairs on a trial was identical for each subject and on no trial did the same digit appear twice. A random order for individual trials was determined in advance for each subject.

The subject was instructed to call out all digits he could remember in any order, as soon as possible after the signal to respond. This signal was delivered 0·8 sec. after the last stimulus pair on a trial. The interval between the end of each pair of stimuli and the onset of the next pair was also 0·8 sec., while the duration of each stimulus pair and the red flash was 0·3 sec.

Results

Table I summarizes results of analysis of variance of percentage of digits recalled correctly. These percentages were transformed, angle = arcsin $\sqrt{\text{percentage correct}}$ (Snedecor, 1956) before analysis. The principal findings were as follows:

(a) Order of presentation of the viewing conditions, i.e. whether digits were projected to nsaal or temporal retinas first, had no significant effect on recall.

(b) There was a significant improvement ($P < 0{\cdot}001$) in recall under the second viewing condition encountered in the experimental session.

(c) Recall of digits projected to the nasal retinas was significantly better ($P < 0{\cdot}01$) than when they were projected to the temporal retinas. Order of efficiency of recall was right nasal (89·4 per cent), left nasal (88·0 per cent), right temporal (87·8 per cent), left temporal (82·4 per cent) retinas.

(d) Information projected to the right eye was recalled significantly better ($P < 0{\cdot}01$) than that projected to the left eye.

TABLE I

Analysis of variance of percentage of digits recalled correctly

Source of variation	MS	*d.f.*	*F*
Groups (A)	4·80	1	—
Subjects within groups	182·44	12	
Total between subjects		13	
Nasal vs. temporal (B)	119·48	1	13·813†
Order of conditions (C)	587·60	1	67·931‡
Left eye vs. right eye (D)	102·60	1	11·861†
A × D	5·41	1	—
B × D	15·23	1	1·761 n.s.
C × D	14·81	1	1·712 n.s.
D × subjects within groups	21·74	12	2·513 n.s.
Residual (error)	8·65	24	
Total within subjects		42	
Total		55	

† Significant beyond the 0·01 level of confidence.
‡ Significant beyond the 0·001 level of confidence.

Discussion

Many short-term memory experiments (e.g. Brown, 1958; Mackworth, 1962, 1963, 1964; Sperling, 1961, 1962) have shown that information to be recalled is available to the subjects for only a limited time. This being so, the more quickly the subject responds, the more he recalls, which suggests that in the present experiment these results were obtained because digits projected to nasal hemi-retinas were processed more quickly (and therefore recalled better) than those projected to temporal hemi-retinas.

Bower and Haley (1964) using a very different technique, arrive at a similar conclusion. In their experiment, they showed that when stimulus durations are very brief in conditions involving binocular rivalry, the rivalry is resolved in favour of nasal hemi-retinas. As stimulus duration is increased to the order of that used in the present experiment, rivalry is resolved in favour of the dominant eye. However, the present findings make it clear that the superiority of the dominant eye is almost wholly attributable to the relatively inefficient recall of digits projected to the left temporal retina.

It is difficult to account for this solely within the context of the experiment. Subjects seem to have taken longer to process information from this region than from any of the other retinal areas. One possible reason for this may be that suggested by Mishkin (1962) to the effect that all visual information from the left hemisphere must be transferred to the right temporal cortex before it is processed. Assuming also that in the optic system, crossed fibres are more effective than uncrossed (Bower, 1966; Bower and Haley, 1964; Crovitz and Lipscombe, 1963;

Sampson, 1962) an advantage would thus be given to input directly to the right hemisphere, and hence to the dominant eye.

This experiment extends the range of conditions in which eye dominance plays a role in recall. Until now, eye differences in recall have been demonstrated only when the stimuli were *perceived* as overlapped when projected binocularly (Sampson and Spong, 1961*a*, *b*; Sampson and Horrocks, 1967). It now appears that the effect may be demonstrable in any perceptual situation which is novel and or difficult.

References

Bower, T. G. R. (1966). A local sign for depth. *Nature, Lond.* **210**, 1081–2.

Bower, T. G. R. and Haley, L. J. (1964). Temporal factors in binocular vision. *Psychon. Sci.* **1**, 409–10.

Brown, J. (1958). Some tests of the decay theory of immediate memory. *Q. Jl exp. Psychol.* **10**, 12–21.

Buxton, C. E. and Crosland, H. R. (1937). The concept of "eye preference." *Am. J. Psychol.* **49**, 458–61.

Crovitz, H. F. and Lipscombe, D. B. (1963). Dominance of the temporal visual fields at short durations of stimulation. *Am. J. Psychol.* **76**, 631–7.

Gillies, S. M., MacSweeney, S. A. and Zangwill, O. L. (1960). A note on some unusual handedness patterns. *Q. Jl exp. Psychol.* **12**, 113–6.

Mackworth, Jane F. (1962). The visual image and the memory trace. *Can. J. Psychol.* **16**, 55–9.

Mackworth, Jane F. (1963). The duration of the visual image. *Can. J. Psychol.* **17**, 63–81.

Mackworth, Jane F. (1964). Interference and decay in very short-term memory. *J. verb. Learn. verb. Behav*, **3**, 300–8.

Mishkin, M. (1962). A possible link between interhemispheric integration in monkeys and cerebral dominance in man. In Mountcastle, V. B. (Ed.), *Interhemispheric Relations and Cerebral Dominance*, pp. 101–11. Baltimore: John Hopkins Press.

Sampson, H. (1962). Binocular interaction in serial adding. *Aus. J. Psychol.* **14**, 9–21.

Sampson, H. and Horrocks, J. B. (1967). Binocular rivalry and immediate memory. *Q. Jl exp. Psychol.* **19**, 224–31.

Sampson, H. and Spong, P. (1961*a*). Handedness, eye-dominance and immediate memory. *Q. Jl exp. Psychol.* **13**, 173–80.

Sampson, H. and Spong, P. (1961*b*). Binocular fixation and immediate memory. *Br. J. Psychol.* **52**, 239–48.

Snedecor, G. W. (1956). *Statistical Methods.* Iowa: Iowa State College Press.

Sperling, G. (1960). The information available in brief visual presentations. *Psychol Monogr.* No. 498, **74**, No. 11.

Sperling, G. (1962). Visual information storage. In *Proceedings of the Sixteenth International Congress of Psychology*, pp. 558–9. Amsterdam: North-Holland.

Received 9 July 1968

Q. Jl exp. Psychol. (1969) **21,** 43–54

DIRECT MEASURES OF SHORT-TERM VISUAL STORAGE

RALPH NORMAN HABER AND L. G. STANDING
Department of Psychology, University of Rochester

Two experiments, involving seven conditions, explored the use of direct measures of visual persistence. In each, the subject was asked to judge if an intermittent stimulus appeared perceptually continuous, or whether it completely faded before the next presentation occurred. The first experiment showed that visual persistence was set at approximately 250 msec. for a recycling presentation of a circle in a tachistoscope; in another task employing a moving opaque slit passing back and forth over a circle, persistence times averaged 50 msec. longer. Reducing luminance by 2 log units increased persistence only slightly, though removing the adapting field increased it by over 100 msec. The second experiment, using the repeating circle, varied the duration of the stimulus, and compared monoptic with dichoptic presentations. Visual persistence was found to be independent of stimulus duration over a range of 4 to 200 msec., where all durations were above recognition threshold for the stimulus. Persistence was unaffected whether the stimulus was repeatedly presented in the same eye or alternated between eyes, strongly suggesting that the storage is central. Finally, a re-analysis of Dodwell and Engel's paper on stereopsis suggests that their effects can be adequately explained by visual persistence of the asynchronous stereo pairs, rather than a more complex fusion model. All of these results strongly support the use of visual persistence as a direct measure of short-term visual storage.

Introduction

Although the "persistence of vision" has been recognized for at least a hundred years (Helmholtz, 1866, reprinted in 1962), interest in the topic has grown since the publication of Sperling's influential monograph (Sperling, 1960). Visual persistence refers to the possibility that the neural representation of a stimulus persists for some time after the offset of the physical energy impinging on the receptor surface. Using the method of partial report, Sperling found that a brief visual stimulus (e.g. 50 msec.) persists perceptually for 300 msec. or more (and considerably longer if a dark post-stimulus field is employed). Keele and Chase (1967), also using partial report, estimate this storage time as 250 msec., while Averbach and Coriell (1961), using both partial report and a metacontrast technique, obtain values of approximately 250 msec. in both cases. Other studies of short-term storage have been made by Mackworth (1963), Eriksen and Collins (1967), and Smith and Carey (1966) with broadly similar results.

The partial report technique is an indirect indicator of persistence. The assumption is made that under total report instructions, the subject is limited in the amount of information he can recall from the stimulus because of processing and memory limitations, and not because he failed to perceive most or all of the content of the stimulus. To eliminate this source of variance, Sperling asked his subjects to report only part of the content, that part designated for report at varying times after the stimulus was presented. He found, as have others, that the indicator could be delayed several hundred msec. and still permit greater accuracy than under full report conditions. If the subject is asked to report all of 12 items, and is correct on only four of them, he has only 25 per cent of the information available at the time of report. If he is asked for only one row of four items, and the subject reported all four, then at the time the indicator arrived, the subject must have had 100 per cent of the content of the stimulus available.

This procedure, and all of its variations, is indirect in that persistence is inferred from the decay in accuracy of report of partial information, under the sampling assumption spelled out above. It is also a very time-consuming experimental procedure requiring many hours of the subjects' time. A more direct procedure is available, though it has not been seriously exploited. The most direct version would be to present the subject with a brief flash from an optical source of very precise rise and decay times, and ask the subject to rate or estimate its apparent duration. If brief stimuli actually persist for several hundred msec. beyond their offsets, such persistence should be measurable in estimates of apparent duration. This procedure is a very unreliable one, however, since the subject has no good standards for determining the absolute values of very brief pulses.

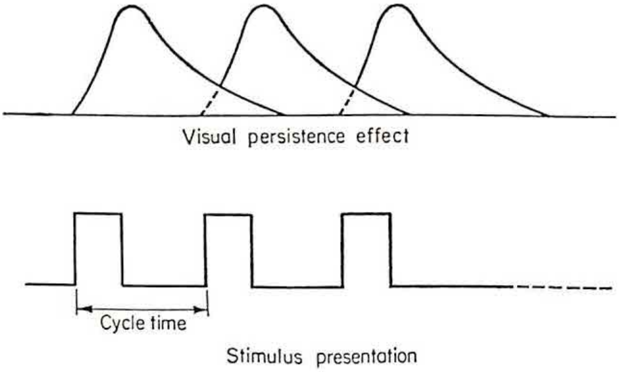


FIGURE 1. A hypothetical view of visual persistence as a result of a repeating stimulus. The subject's response of continuity or discontinuity is based upon whether he judges the persistence reached zero from the first flash before the second flash occurred.

Sperling (1967) reports a variation of it, although he has not used it to determine persistence. At the same time as the subject is given a brief visual flash, he hears two auditory clicks, and is told to adjust the interval between the clicks to match the apparent duration of the visual pulse. Under conditions in which little or no visual persistence is to be expected (visual noise follows the offset of the stimulus immediately) the click interval is set at just slightly longer than the actual stimulus duration. Unfortunately, Sperling does not report data for conditions without visual noise in which visual persistence would be expected.

A second technique, several variations of which are used in the experiments to be described here, presents the stimulus repeatedly with specific delays between each repetition. The subject is asked to judge whether the perception of the stimulus had completely faded before the next flash occurred. The time interval can then be varied in order to locate the threshold of persistence.

Two experiments are reported; in the first a comparison is made between two tasks designed to provide direct measures of storage times, while the second investigates three paradigms to measure storage times under monoptic and dichoptic conditions. The intention of the second experiment is to relate the concept of storage to the findings of Dodwell and Engel (1963) who have demonstrated the production of stereopsis from pairs of monocular stimuli that are presented alternately, and do not overlap in time. It is hoped that these experiments will demonstrate the feasibility of direct measures of storage times, and demonstrate how these measures may be applied to a problem in form perception.

Experiment I

The two tasks employed in this study are both intended to provide a measure of storage time, but are otherwise quite different, since Task 1 uses a moving-slit apparatus, while Task 2 employs repetitive stimuli in a tachistoscope. Task 1 has been described previously by Haber and Nathanson (1968), in more general terms, and is based on one used by Parks (1965). Briefly a narrow vertical slit oscillates, in a horizontal plane, immediately in front of a stationary black-on-white form. The subject views the apparatus with steady fixation. When the rate of oscillation is sufficiently high, a whole form is perceived even though the slit exposes only a small portion of the form at any one instant. As the frequency of oscillation is increased from zero, Haber and Nathanson (1968) have reliably observed the following stages: the shape of the form may be inferred, but the form is not *perceived* in the true sense of the word; the form is perceived as a whole, but it does not persist between successive sweeps of the slit, so that it is seen as present intermittently; the form is seen as both whole and as continuous in time. We utilize the beginning of the third stage as indicating the point at which short-term storage time is exactly equalled by the sweep time of the slit.

Task 2 uses a brief trachistoscopic presentation of the whole test form, followed by a long exposure of a blank adapting field at the same luminance as the test field background. This cycle is repeated continuously, and the frequency is varied to find the frequency at which the form appears to be present continuously, rather than appearing intermittently. Again, the cycle time at this point provides a measure of how long the short-term storage must last for this phenomenally continuous percept to occur. As in the first task, we assume that the subject is capable of making meaningful "psychological reports" (Natsoulas, 1967) concerning his percepts. Figure 1 presents a hypothetical view of the subjects' report in Task 2.

Method

Subjects. Twenty-four experimentally naïve subjects were employed. Nine were graduate students in psychology at the University of Rochester; most of the remainder were non-psychology undergraduates. All possessed normal or corrected vision. Two potential subjects were rejected due to inability to obtain the "illusion" which is the basis of Task 1.

Apparatus and experimental tasks. The apparatus for Task 1 has been described by Haber and Nathanson (1968). It consists of a vertical oscillating slit driven by an electric motor, the rate of oscillation being precisely controlled by the experimenter by means of a continuously variable speed control on the motor. The slit, cut in a large piece of white cardboard, was 5 cm. high and 0·2 cm. wide; it oscillated horizontally through a sweep of 4 cm. The stationary test stimulus was located 0·1 cm. behind the slit, at the midpoint of the sweep, and comprised a circle, 3 cm. in diameter, drawn in black with a 0·1 cm. line, upon a white card. The contrast ratio was approximately 90 per cent. Due to its size, only a small portion of the circle was exposed at each instant in time as the slit swept back and forth across the circle; the mean proportion of the time for which each element of the circle was exposed to the subject was 1 in 20. The apparatus was viewed from a distance of 30 in. and was illuminated from directly in front. Fixation upon the centre of the circle may be maintained quite easily under these conditions.

The apparatus used in Task 2 was a Scientific Prototype two-channel tachistoscope (model 800F). A circle, identical to that used in Task 1, was exposed in the test channel, followed by an exposure of a blank adapting field, at the same luminance, in the second channel; this sequence was repeated continuously. The ratio of the exposure durations for the stimulus and the blank channel was always maintained at 1 to 20, while the length of the cycle (test plus blank field duration) was varied to find the threshold for apparent visual persistence for the form. The length of cycle time at this point provides a measure of the short-term storage time required to produce this phenomenal effect. As in Task 1, the viewing distance was 30 in. Durations of the tachistoscope were calibrated using a photomultiplier tube and oscilloscope. Calibration is necessary for extreme duty cycles (on to off ratio of more than 1 to 10) with this tach.

Procedure. Two versions of each task were used: high and low luminance (0·5 and 50 ml., respectively). Each subject was tested with each of the four resultant task-luminance combinations; the subjects were arranged in six 4 × 4 Latin squares which were formed using all 24 possible orders of the four experimental treatments. All four treatments were given in a single experimental session of about 2 hr.; 10 threshold measurements were taken in each, using alternate ascending and descending series in the method of limits. Steps of 20 msec. in the cycle time were used.

Measurements of storage time were also made in Task 2 with 2 of the subjects after the main experiment was finished under conditions where a totally dark adapting-field alternated with the test stimulus (at 50 ml.); all other conditions were as before.

For Task 1, the subject was instructed, "While the experimenter adjusts the speed of the moving slit, maintain fication steadily at the mid-point of its movement. You may be able to see a circle behind the slit; on each trial, which will be indicated by the experimenter saying 'now,' say 'C' if the circle appears to be present and continuous in time, 'D' if it appears to be present but discontinuous in time, or 'zero' if you do not see the circle. The perceived continuity of the circle need not be perfect: it may appear to fluctuate somewhat without losing the appearance of being present continuously. However, you should rate the circle as discontinuous if you can perceive a blank interval, however brief, between presentations of the circle. Do not allow the presence or absence of apparent movement of the circle, or of flicker in the background, to influence your report of what you perceive, as these effects are irrelevant to the phenomenon under study."

For Task 2, the subject was instructed, "Fixate the circle, which will be presented in the tachistoscope, and say 'C' if the circle appears to be present continuously, 'D' if it appears to be discontinuous, or 'zero' if you do not see the circle." The remaining instructions were identical to those for Task 1.

The subjects were encouraged to ask questions if they did not fully understand the task. While they were kept unaware of specific experimental hypotheses, they were always informed that even if the test stimuli sometimes appeared continuous in time, they were not physically so. Thus, while subjects had a specific set to look for continuity or visual persistence, they were aware that perceived continuity was always illusory. It was stressed that there were no "right" or "wrong" judgements, and that a direct report of the subject's perceptual experience was desired.

Results

The ratings of "C" and "D" given in each task for the ascending and descending presentations were used to calculate each subject's persistence threshold for storage time under each treatment. No "zero" responses occurred. Reversals of ratings occurred in a few series (about 15 per cent), and were usually confined to two or three trials. The ten threshold values were averaged for each subject-treatment combination, to yield four mean storage times for each subject (see Table I).

TABLE I

Mean storage times for Task 1 (*moving-slit apparatus*) *and Task* 2 (*tachistoscope*), *under high and low luminance conditions, and for two subjects in Task* 2 *in which the adapting field was off*

Task	1		2		·
Luminance level (ml.)	50	0·5	50	0·5	50 (dark interval)
Mean storage time (msec.)	319·1	338·9	261·2	295·7	405·0
Standard deviation (between subjects) (msec.)	33·5	31·6	25·7	29·3	37

The ten thresholds on which each mean is based were found to have a standard deviation of 31·2 msec., within subject-treatment combinations. To facilitate comparison between Task 1 and Task 2, storage times in each task were taken as the total time at persistence threshold, i.e. the sum of stimulus and blank field exposure durations, rather than blank field duration alone. (If storage time were expressed as blank field duration alone, this would reduce the calculated storage times by only 5 per cent.)

The mean storage times obtained above were submitted to an analysis of variance. An identical analysis was also performed on the standard deviations of these storage times. The first analysis showed that the storage times obtained in Task 1 were significantly longer than those obtained in Task 2 ($F = 11{\cdot}5$, 1 and 15 *d.f.*, $P < 0{\cdot}01$). For both tasks, storage times were longer under low, as compared to high, liminance, but this difference failed to reach statistical significance ($F = 3{\cdot}14$, 1 and 15 *d.f.*, $P > 0{\cdot}05$). However, a sign test of this difference indicated that it was significant at the 0·05 level (32 positive differences in 48 comparisons, $P = 0{\cdot}03$). The interaction between luminance level and tasks was not significant ($F = 2{\cdot}1$, 1 and 15 *d.f.*, $P > 0{\cdot}05$).

The first analysis also showed that subjects differed significantly in their mean storage times ($F = 8{\cdot}5$, 23 and 15 *d.f.*, $P < 0{\cdot}01$). While this factor is confounded with order of testing, the latter is shown to be irrelevant, since no differences were found over the successive stages of the four experimental treatments for each subject ($F < 1$), and individual differences were equally pronounced on the first task performed by each subject. Pooling each subject's storage times for high and low luminance levels, the correlation (rho) between storage times on the two tasks was found to be 0·37, a value which is significantly greater than zero ($P < 0{\cdot}05$).

The second analysis, based on the standard deviation of the ten threshold estimates for each subject-treatment combination, showed that variability of performance was not significantly affected by any factor except subjects ($F = 6{\cdot}6$, 23 and 15 *d.f.*, $P < 0{\cdot}01$).

The mean storage time for dark adapting-field conditions (see last column on Table I) was found to be 390 and 420 msec., for two subjects, with a combined standard deviation of 37 msec. (The mean storage times for these subjects were 273 and 240 msec. when the test and adapting fields were both at 50 ml.)

Discussion

The results show that in four experimental situations subjects are able to make judgements of apparent visual persistence, with a standard deviation of about 10 per cent of their mean values. Thus, direct measures of storage time are at least capable of providing reliable data, despite the use of inexperienced subjects. No difficulties were found with this technique, it yields values which are consistent with other approaches to the study of short-term storage, and its simplicity, speed and reliability suggest it is a useful method for examining the factors which determine the efficacy of visual storage mechanisms. For example, we have readily found increased storage times to occur when a dark interval is substituted for the lighted adapting field in Task 2, as would be expected from Sperling's (1960) results.

Regarding the validity of the data, the results indicate that on average a form presented for approximately 15 msec. persists perceptually for about 300 msec. This agrees well with the estimates of Sperling (1960) and of Averbach and Corriell (1961) for similar, lighted adapting field conditions: 300 and 250 msec., respectively. While test stimulus durations are longer in these studies than in the present case, Sperling (1960) has suggested that for above-threshold stimuli, storage time is almost independent of stimulus duration, so that these studies may meaningfully be compared with the present findings. (We also have found—see Experiment II—the storage time to be largely invariant under wide changes in stimulus duration.) Thus we believe that the two tasks employed in the present study do in fact measure the visual persistence of brief stimuli in much the same way as do the more elaborate but cumbersome partial report and metacontrast methods.

The storage effect is very robust under changes in most stimulus parameters (except for adapting-field luminance, according to the results of the dark-condition data on two subjects), and this is illustrated by the present finding that a reduction in test stimulus luminance of 2·0 log units hardly alters the storage time. The tendency for storage time to increase by 20 to 30 msec. for the low-luminance condition (which is about 0·5 to 0·7 log unit above threshold for recognition of the circle), if this is a genuine effect, may result from the fact that the adapting field is also of lower luminance and thus less likely to obliterate the image of the circle. A lower level of light adaptation is also known to reduce the temporal sensitivity of the eye appreciably, by increasing the critical duration for visual integration of luminance (Barlow, 1958), and by lengthening the perceptual latency of the visual system in some cases (Standing, Dodwell and Lang, 1968). These effects would presumably reduce the ability of the eye to detect the discontinuity of the intermittent test stimulus in the present study.

A more surprising feature of the results is that storage time as measured by Task 1 was approximately 50 msec. longer than that measured by Task 2, although both estimates are plausible in view of other studies of short-term storage. While the optical quality of the image in Task 1 is inevitably lower than that in Task 2, and the subject must refrain from making eye movements if the image is not to vanish altogether, the moving-slit apparatus does eliminate the slight but detectable transient in field luminance which invariably occurs when switching fields in a tachistoscope. It may be that this transient acts to a slight extent as a "backward masking" stimulus (Raab, 1963) to reduce the adequacy of the image stored in the visual system. Various other possibilities exist: subjects may adopt a lower criterion for visual persistence in the first task, or the visual system may respond more slowly to a form that is presented a small section at a time, rather than all at once (with repetition rate held constant). This would be analogous to the finding in flicker-fusion studies that flicker is harder to perceive when a dim light, or one that is spread out over time, is employed (Bartley, 1963).

While it would be valuable to determine the reason for the longer storage times in Task 1, the finding of fairly similar values for Task 2 goes some way to confirm the analysis made by Haber and Nathanson (1968) of Parks' (1965) results. The "camel seen through the eye of a needle" effect is not based on post-retinal re-integration of visual forms, as Parks suggested, but on the "painting" of these forms out across the retina, and their slow decay in the short-term store, so that the next presentation of the stimulus occurs when there is still an appreciable trace in the visual system from the previous presentation. However, although the moving-slit task is easier for subjects to judge, according to their reports, Task 2 is preferable as a technique since it provides much greater flexibility in varying parameters.

The differences found between subjects in storage time are no larger than individual differences typically found in psychophysical studies. Sperling (1960) has shown that the stimulus trace decays in a gradual, monotonic curve; therefore subject in a very real sense must adopt a criterion, to be able to specify the inter-stimulus interval at which the trace just decays to zero before the next presentation of the stimulus. The reliability of judgements in this study (and post-experimental reports by the subjects) suggest that they are able to impose this cut-off on their perceptual experience without excessive difficulty.

The low (but significant) correlation of storage times between the two tasks suggests that the tasks were sufficiently different so that subjects reacted to them in somewhat different ways. Until more attention is paid to individual differences in psychophysical tasks, further interpretation of the magnitude of this correlation is highly speculative.

Experiment II

One aim of this experiment is to study visual storage time as influenced by duration of the test stimulus, which Sperling (1960) suggests should not appreciably alter the effect. Another aim is to relate the results from three versions of tachistoscopic task employed in Experiment 1 to measures of "stereopsis time" obtained in studies where the two eyes are stimulated, asynchronously, with the two members of a stereo pair (Dodwell and Engel, 1963; Efron, 1957).

Three paradigms were employed in this study; they are illustrated in Figure 2, together with the paradigm used by Dodwell and Engel (1963) to investigate binocular fusion. In the latter case, the dependent variable was the appearance of stereo depth in a random brightness field (Julesz, 1965), such that the depth was perceptually continuous, whereas the dependent variable in the present paradigms is the appearance of perceived continuity of non-stereo stimuli. In both cases, it is convenient to use Dodwell and Engel's notation of SI and LI, respectively, to refer to the "short interval" within pulse pairs (or the duration of a single pulse, as in paradigm 1), and the "long interval" between each pair of pulses (or, the inter-pulse interval, as in paradigm 1).

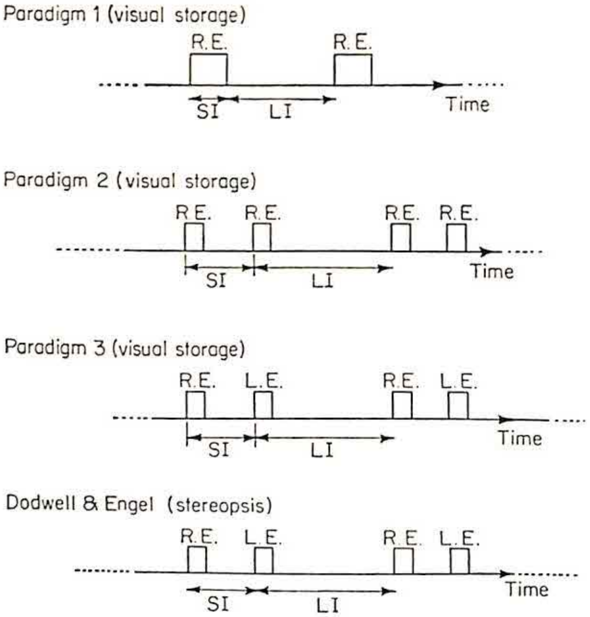

Paradigm 1 (visual storage)
R.E.
R.E.
SI
LI
Time
Paradigm 2 (visual storage)
R.E.
R.E.
R.E.
R.E.
SI
LI
Time
Paradigm 3 (visual storage)
R.E.
L.E.
R.E.
L.E.
SI
LI
Time
Dodwell & Engel (stereopsis)
R.E.
L.E.
R.E.
L.E.
SI
LI
Time

FIGURE 2. Presentation paradigms used in Experiment 2. The abscissa represents the presentation of the test stimulus to the right eye (R.E.) and left eye (L.E.), respectively. An adapting field (not indicated) is present between each pulse of the test stimulus.

Method

Subjects. Three new subjects were used, one each in paradigms 1, 2 and 3. All were experimentally naïve and possessed normal uncorrected vision.

Apparatus and tasks. A three-channel Scientific Prototype tachistoscope (model GA) was used. In paradigm 1, single repeated exposures of the test stimulus (a circle, identical to that used in Experiment 1), are presented monocularly in one channel of the tachistoscope for SI msec., followed by a blank adapting field for LI msec. in the second channel. In paradigm 2, the stimulus is presented monocularly in the first channel, as in paradigm 1, except that two 10 msec. pusles separated by SI msec. are used. In paradigm 3, two 10 msec. pulses separated by SI msec. are presented to the right and left eye, respectively, using two channels of the tachistoscope and an arrangement of Polaroid filters. In paradigms 2 and 3, the interval, LI msec., between pairs of pulses is filled with an adapting field. Test and adapting field luminances were set at 20 and 10 ml., respectively, for all paradigms.

Procedure. Using the pardigms described above, the procedure followed that of Task 2, Experiment 1. The threshold for perceived continuity of the test stimulus, at each of 12 SI durations ranging from 4 to 200 msec., was determined by varying LI in steps of 20 msec., and obtaining the subject's judgement of continuity versus discontinuity at each LI value. Threshold is taken as the LI value which produces 50 per cent "continuous" responses. Two ascending and two descending series were made of each threshold, over a range of 8 LI steps.

Results

The data for the three paradigms are displayed in Figure 3, in which changes in persistence (LI) threshold are plotted against log exposure duration (SI). It is clear that for each paradigm, persistence is invariant over a wide range of stimuli durations. Further, each of the three procedures yields approximately the same storage time (220 msec.), even with different methods of presenting stimuli.

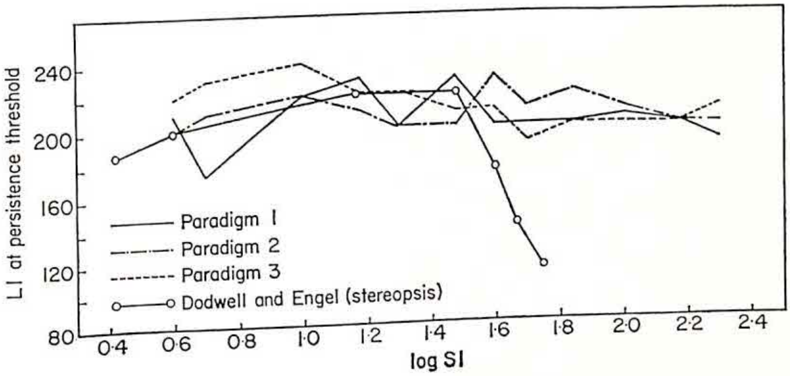


FIGURE 3. Persistence threshold (maximal LI for which subject judges continuity of persistence) as a function of log SI, for each of the three paradigms of Experiment II. Each curve is for one subject, with each point based upon four threshold determinations. Also: threshold values of LI for stereopsis, redrawn from data of Dodwell and Engel.

Discussion

Experiment 2 yields several important results. That visual persistence is independent of exposure duration supports a basic argument of Sperling (1960, 1967) regarding the independence of the duration of a short-term visual storage from the duration of the stimulus that created it. Sperling's evidence rested primarily on indirect measurement of storage—the probability of reporting part of the stimulus array correctly as a function of delay of the indicator, varied over a range of 10 to 500 msec. Recently, Raymond and Glanzer (1967) reported an apparent failure to replicate this invariance under conditions of presence and absence of either a pre- or a post-visual noise field. However, they use very low luminance levels (1·5 log unit lower than Sperling's), so it is quite likely that the stimulus array was at or below threshold for the briefest flashes. There is no reason to believe that persistence of storage should be as long for a subthreshold flash as for a supra threshold presentation, and certainly, there is nothing in Sperling's work to suggest this prediction.

Paradigm 3 provides a test of a central vs. peripheral locus of visual storage, and the results strongly support a central interpretation. Using monocular presentations (paradigm 2), the stimulus must be repeated in the same eye within 220 msec. for the subject to judge it as continuous, and any delay beyond 220 msec. elicits a response of discontinuous. However, using dichoptic presentations, a delay as long as 440 msec. for either eye will still produce continuity, as long as the other eye also received a flash halfway through that time interval. Hence, it is not the individual receptor that must be stimulated at least as often as every 220 msec., but some central locus that is combining inputs from the two receptors. This finding does not eliminate the psosibility that the neural persistence may be retinal, in the sense that a brief (but above threshold stimulus flash) may continue to produce excitation for several hundred msec. which is transmitted to the brain. However, the subject's judgement of persistence (and in Sperling's work, his processing of information) must be based upon some central events, and not simply upon persistence of retinal stimulation. This argument, among other things, makes the after-image an improbable mechanism for visual persistence underlying short-term visual storage.

Figure 3 shows that the short-term visual storage effect lasts at least as long as the intervals over which Dodwell and Engel (1963) were able to obtain stereopsis. It seems that their results, while empirically correct, reflect storage processes and not mechanisms of binouclar fusion as these authors believe. The assumption by Dodwell and Engel that they have demonstrated the quantitative features of a central mechanism for stereopsis appear unwarranted. Our view is that, whatever the mechanism for perception of depth as a result of disparity between two dichoptic stimuli, all that is required is that the two inputs be perceptually present at the same time, a point which was clearly recognized by Munsterberg (1896). The present study suggests that this condition may easily be fulfilled, due to the storage effect, at interstimulus intervals where Dodwell and Engel have found stereopsis to occur. We also note that Kahneman, Norman and Kubovy (1967) have found perfect binocular summation in an acuity task when two successive flashes of 160 msec. each are presented dichoptically. Again, sizeable dichoptic temporal integration must be occurring even though no special mechanism for stereopsis can be involved.

Julesz (1965) has produced evidence, from a masking study, that the time constants for the perception of depth are longer than for the perception of form in the individual stereo stimuli. This finding is to be expected: the temporal sensitivity of a higher-order process (stereopsis) can only be expected to fall below that of the constituent processes (form perception). It also tends to support our contention that the threshold curve of Dodwell and Engel indicates the extent of visual storage (under the rather different conditions of their study), rather than the limits of the central nervous system's ability to integrate asynchronous stereo information to infer depth, since it suggests that failures in perceived continuity will occur first in the perception of the monoptic forms, as LI is increased, while the sensation of depth can persist at higher values of LI. Thus, we infer, the threshold value of LI will not reflect the temporal properties of the mechanism for stereopsis.

Conclusions

The two experiments taken together support the following statements and interpretations.

(1) Visual persistence can be directly measured by asking subjects to judge when a repeating stimulus has completely faded before its next appearance. These judgements are stable within and between subjects.

(2) This stability of persistence was assessed in two somewhat dissimilar tasks with remarkably similar overall results.

(3) The thresholds for persistence were on the same order of magnitude of those for the indirect assessment of the duration of short-term visual storage (250 to 350 msec.) so that it seems likely that measurement of visual persistence is a direct assessment of short-term visual storage.

(4) Visual persistence is slightly longer (20 to 30 msec.) when luminance is reduced by 2 log units. It is markedly longer (120 to 180 msec.) when no adapting and background field is used. Both of these findings agree both with data on short-term visual storage, and the general effect of luminance changes in the eye.

(5) Visual persistence is independent of the exposure duration of the stimulus over a range of 4 to 200 msec. This supports Sperling's findings for above threshold stimuli.

(6) Visual persistence is the same whether the repeating stimulation comes in the same eye, or alternates between eyes. This strongly supports a central locus for the effects of persistence and of short-term visual storage.

(7) Finally, Dodwell and Engel's work on binocular fusion and the perception of depth can be much more parsimoniously explained as the persistence and decay of dichoptically presented stimuli, rather than through some special properties of stereopsis.

This research was partially supported by grants from the United States Public Health Service MH 10753 and from the National Science Foundation GB 5910 to the first author. We would like to thank Miss Barbara Herr for her assistance in some of the data collection.

References

Averbach, E. and Coriell, A. S. (1961). Short-term memory in vision. *Bell Syst. tech. J.* **40**, 309–28.

Barlow, H. B. (1958). Temporal and spatial summation in human vision at different background intensities. *J. Physiol., Lond.* **141**, 337–50.

Bartley, S. H. (1963). *Vision: A Study of its Basis.* New York: Hafner.

Dodwell, P. C. and Engel, G. R. (1963). A theory of binocular fusion. *Nature, Lond.* **198**, 4875, 39–40 and 73–4.

Efron, R. (1957). Stereoscopic vision. I. Theory of binocular temporal summation. *Br. J. Ophthal.* **61**, 709–30.

Eriksen, C. W. and Collins, J. F. (1967). Some temporal characteristics of visual pattern recognition. *J. exp. Psychol.* **74**, 476–84.

HABER, R. N. and NATHANSON, L. S. (1968). ' Post-retinal storage? Some further observations on Parks' Camel as Seen through the Eye of a Needle. *Percept. Psychophys.* **3**, 349-55.

HELMHOLTZ, H. von. (1962). *Physiological Optics*, Vol. III. New York: Dover (republished), 205–6.

JULESZ, B. (1965). Texture and visual perception. *Scient. Am.* **212**, 38–48.

KAHNEMAN, D., NORMAN, J. and KUBOVY, M. (1967). Critical duration for the resolution of form: centrally or peripherally determined? *J. exp. Psychol.* **73**, 323–7.

KEELE, S. W. and CHASE, W. G. (1967). Short-term visual storage. *Percept. Psychophys.* **2**, 383–6.

MACKWORTH, J. (1963). The duration of the visual image. *Can. J. Psychol.* **17**, 62–81.

MUNSTERBERG, H. (1896). A stereoscope without mirrors or prisms. *Psychol. Rev.* **1** 56–60.

NATSOULAS, T. (1967). What are psychological reports about? *Psychol. Bull.* **67**, 249–72.

PARKS, T. E. (1965). Post-retinal visual storage. *Am. J. Psychol.* **78**, 145–7.

RAAB, D. (1963). Backward masking . *Psychol. Bull.* **60**, 118–29.

RAYMOND, B. and GLANZER, M. (1967). Continuity of exposure length effects in tachistoscopic perception. *Psychon. Sci.* **7**, 287–8.

SMITH, F. and CAREY, P. (1966). Temporal factors in visual information processing. *Can. J. Psychol.* **20**, 337–42.

SPERLING, G. (1960). The information available in brief visual presentations. *Psychol. Monogr.* **74**, 11 (Whole no. 498).

SPERLING, G. (1967). Successive approximations to a model for short-term memory. *Acta Psychol.* **27**, 285–92.

STANDING, L. G., DODWELL, P. C. and LANG, D. (1968). Dark-adaptation and the Pulfrich effect. *Percept. Psychophys.* **4**, 118-20.

Received 8 *August* 1968

Q. Jl exp. Psychol. (1969) **21,** 55–57

STIMULUS CHANGE EFFECTS OF FORCING ON MAZE ARM SELECTION IN HAMSTERS

R. N. HUGHES† AND D. E. BLACKMAN‡
Department of Psychology, Queen's University, Belfast

Hamsters were forced to enter one arm of a T-maze on four succesive runs by means of either a grey door of the same brightness as the rest of the apparatus or by a black or white door. During four successive free choices the subjects entered the previously blocked arm significantly more often following forcing with the black or white door.

Introduction

Dember and Fowler (1959) have shown that "forcing" a rat to enter one arm of a T-maze by blocking the alternative will tend to increase alternation on a following "free" trial without any forcing barrier present. Dember (1961) suggested that this effect could be accounted for by increased attractiveness of the previously blocked alternate arm due to the stimulus change produced by intertrial removal of the forcing barrier. Tentative evidence for this view was produced by Hughes (1966) who, using rats, found more initial orientation towards and partial entries of the alternate arm of a T-maze following forcing with a wooden barrier than following either no forcing or forcing with a "less perceptible" transparent barrier. (It was argued that intertrial removal of the transparent barrier represented the lowest degree of change.) The present study was designed to retest Dember's explanation. It was felt that removal of a black or white barrier from a mid-grey T-maze would represent a higher degree of change than removal of a mid-grey barrier from the same apparatus. As Dember and Millbrook (1956) have shown that arm preferences are directly related to the amount of intertrial brightness change, it was hypothesized that there would be a greater tendency to choose the arm previously blocked with a black or white forcing barrier in the course of four successive free trials.

Method

Subjects and apparatus

The subjects were 12 Golden Hamsters (8 male, 4 female) about 7 months old, kept in individual cages in the experimental room. The apparatus was a wooden T-maze with walls 6·5 in. high and a startbox 10 in. long. The stem was 24 in. long and the two arms were each 18 in. long. At the end of each arm there was a 10-in. long goal box of the same width and height as the arm with which it was connected by a 4 × 4 in. opening. The width of the maze, startbox and goalboxes was 6 in. throughout. The whole apparatus was covered by hinged, wire-gauze lids and its interior was painted mid-grey. The arms could be blocked off from the stem by painted, aluminium guillotine doors.

---

† Now at Department of Psychology, University of Canterbury, New Zealand.
‡ Now at Department of Psychology, University of Nottingham.

Procedure

The general procedure for each subject involved a series of four successive forced trials followed by four successive free choices for eight consecutive days. There were two forcing conditions. In condition G the subject was forced left or right four times by means of a barrier of the same brightness as the rest of the maze, i.e. mid-grey. In condition B/W, it was forced four times with a black (for half the series) or a white barrier (for the other half). In both conditions each forced run was followed by 30 sec. confinement to the entered arm, and then by 30 sec. back in the startbox. At the end of each series of forced trials, the subjects were given four successive free choices being confined to the chosen arm and then the startbox for 30 sec. at the end of each choice. All subjects experienced both conditions four times and were forced left for half their forcing series and right for the other half. The first arm entered and total entries of the alternate, previously blocked arm were recorded for each subject during each series of free choices.

Results and Discussion

The total number of first choice alternations in the four series of free choices in each condition were counted for individual subjects. Also, for each free-choice series in the two conditions every hamster's total entries of the previously blocked arm was determined. From these latter totals, representative individual scores were arrived at by calculating medians of the four free-choice series in each condition for each animal. All scores were then compared between conditions by means of Wilcoxon matched-pairs signed-ranks tests (see Table I).

TABLE I

Medians of the two-response measures following two-change conditions, and results of Wilcoxon tests

Measure	Change conditions		T	N	z
	G	B/W			
Total first-choice alternations	3·00	3·00	6·0	6	0·42
Median entries	2·75	3·00	10·5	10	1·73*

†$P<0·05$.

On the first runs of the four free choice series 10 out of 12 hamsters in condition G and 12 out of 12 in condition B/W chose the previously blocked arm more than 50 per cent of the time. According to two-tailed binomial tests these numbers were significantly greater than chance expectancies (G, $P=0·038$; B/W, $P<0·002$). Likewise, the median entries of the previously blocked arm were greater than chance for 10 and 12 subjects in conditions G and B/W, respectively.

Although there was a significant effect in the direction predicted on *median entries*, this effect would seem to be slight in view of the lack of a significant difference between the two conditions on *total first choice alternations*. Nevertheless, it can be concluded that the hamsters showed on successive free choices a greater preference for entering the arm that had been previously blocked by a black or white door, i.e. the arm in which the greater brightness change had resulted from intertrial removal of the door. This would appear to support Dember's (1961) explanation for the effects of forcing. However, it is also possible that the greater preference for the previously blocked arm in the B/W condition was merely a carry over from the forced to the free trials of preference for the barrier and hence arm that was different in brightness from the rest of the apparatus. Further experimentation could determine which of these two interpretations is the case.

The authors are grateful to Mr. Trevor Williamson who ran the animals.

References

DEMBER, W. N. (1961). Alternation behavior. In FISKE, D. W. and MADDI, S. R. (Eds.), *Functions of Varied Experience*, pp. 227–52. Homewood, Ill.: Dorsey.

DEMBER, W. N. and FOWLER, H. (1959). Spontaneous alternation after free and forced trials. *Can. J. Psychol.* **13**, 151–4.

DEMBER, W. N. and MILLBROOK, B. A. (1956). Free-choice by the rat of the greater of two brightness changes. *Psychol. Rep.* **2**, 465–7.

HUGHES, R. N. (1966). Stimulus-change effects of forcing in alternation behavior. *Psychol. Rep.* **19**, 515–8.

Received 1 *June* 1968

Q. Jl exp. Psychol. (1969) **21**, 58–62

REVERSAL LEARNING OF OBJECT AND POSITIONAL DISCRIMINATIONS BY MINK, FERRETS AND SKUNKS

BARBARA A. DOTY AND WALTER C. COMBS

Department of Psychology, North Central College

Mink, ferrets and skunks were tested on successive reversals of two-choice object or positional discriminations in a modified Wisconsin General Training Apparatus. All species showed considerable improvement in performance after considerable training and some subjects exhibited one-trial reversal learning. Mink and ferrets persistently made more errors when position cues were relevant than when object cues were relevant; the reverse was true of skunks.

Introduction

Improvement in learning repeated reversals of the same discrimination has been observed in a number of mammals as well as in birds and invertebrates (Harlow, 1949; Dufort, Guttman and Kimble, 1954; Cronholm, Warren and Hara, 1960; Warren and Warren, 1962; Kirk and Bitterman, 1963). Performance of primates on repeated reversals of non-spatial discriminations does not differ consistently from that of non-primate mammals or birds (Schusterman, 1962; Warren, 1965; Rollin, 1963; Hughes, 1964; Bitterman, 1963; Bacon, Warren and Schein, 1962). All these groups made progressively fewer errors on repeated reversal trials and they ultimately learned to reverse their responses very rapidly.

In the present study, repeated reversal learning of a single discrimination was observed in certain Mustelidae (weasel) family subjects, namely mink, ferrets and skunks. These species have not been systematically studied under laboratory learning conditions. Performance on spatial and non-spatial visual discriminations was compared to determine whether weasel-family subjects respond differentially to the two types of cues involved. Previous data for cats and monkeys indicate that cats learn spatial discriminations more rapidly than object-quality discriminations, while the reverse is true for rhesus monkeys (Harlow, 1951; Warren, 1965).

Method

The subjects consisted of eight animals of each species of mink, ferrets and skunks. About half the subjects in each species group were males. All animals were Carnivora order, Mustelidae family. Descented skunks were *Mephitis* species, or striped variety, commonly called Illinois skunks, and weighed 12 to 14 lb. prior to this experiment. Ferrets were a commercially developed species, *Mustela furo*, were brown or white in colour, and weighed

7 to 10 lb. Mink were also a commercially developed species; the Mustela vision, pearl variety, and ranged in weight from 6 to 8 lb. Animals ranged in age from 10 to 12 months.

Subjects were experimentally naive and were raised from infancy in the College animal colony. They were maintained on canned cat and dog food supplemented with vitamins, brewer's yeast and fish flour. Animals were housed in 17 × 36 × 17 in. wire cages. Whenever possible, subjects were handled or were permitted to move freely around the College laboratory. These procedures resulted in considerable taming of most animals except mink. The latter showed no tendency toward docility and were transported to the learning apparatus in a commercial self-trapping cage. Other animals were carried by the experimenter.

The apparatus was a modified Wisconsin General Test Apparatus (Harlow, 1949) consisting of two unpainted wooden chambers completely enclosed within a ½ in. plywood container. One chamber, the start box, was 20 × 10 × 20 in. overall with the rear wall serving as an entrance. A 20 × 10 × 20 in. test chamber was separated from the start box by a 20 × 20 in. wall. Centred in the wall 5 in. above the floor was a 10 × 4 in. sliding door which permitted the subject access to the test chamber. A 20 × 7 in. one-way mirror through which the experimenter viewed the subject was located on the outside test chamber wall opposite the sliding door. The test chamber was illuminated by a 6 W. fluorescent light placed in the ceiling over the mirror. Below the mirror was a 20 × 20 in. door through which a sliding tray was passed by the experimenter. The tray served as a platform to support two 3 × 6 × 3 in. alleys, each equipped with a 3 × 3 in. door, painted in flat black, hinged at the top and opening inward to each alley.

The test objects consisted of a triangle, the apex of which was oriented vertically, and a square. The objects were wooden, and were 2 sq. in. in area and 1½ in. thick. Objects were painted white and were fastened to the alley doors to prevent subjects from removing them.

Initial discrimination training

One-half of subjects within each species group were randomly assigned to a positional or to an object discrimination problem. For one-half the subjects in the object discrimination group, choice of the triangle was always rewarded; for remaining animals, choice of the square was reinforced. The right-left position of the rewarded object was varied in a balanced, irregular sequence. One-half of subjects in the position discrimination group were required to select the right alley door; remaining subjects had to select the left alley door to obtain the food reward which was a ½ in. cube of meat located in the alley behind the correct door. Spatial position of objects was varied from trial to trial in a balanced, irregular sequence. Thus, irrelevant object cues were present on the positional discrimination. Training was continued until the subject attained a criterion of 10 consecutive correct responses.

Reversal training

After the subject met this criterion, reversal trials were begun and the subject was required to select the previously incorrect object or door. Subsequently the two objects or doors became alternately correct and incorrect as the above criterion was met. The non-correction method was used throughout testing.

Animals were trained on a series of 60 reversals. They performed 1 reversal a day, 7 days a week. Subjects were returned to their home cages after the last trial on each day and were immediately fed daily rations. Animals were maintained at 80 per cent of normal body weight.

Results

Initial discrimination learning

Median errors to criterion in initial discrimination learning for subjects trained on object or positional discriminations are presented in Table I. Mink and ferrets

TABLE I

Median errors in initial discrimination learning by mink, ferrets and skunks

Subjects	Discrimination Object	 Position	U
Mink	7·5	21·2	0†
Ferrets	10·3	25·4	1‡
Skunks	36·7	10·4	0

† $P < 0·028$, two-tailed test.
‡ $P < 0·058$, two-tailed test.

made significantly fewer errors on the object discrimination while the reverse was true of skunks.

Reversal learning

Median errors to criterion, plotted against successive blocks of ten reversals are summarized for each species separately in Figure 1. Errors to criterion on the first and last ten reversals for subjects of each species attempting each type of discrimination were analysed by Wilcoxon matched-pairs signed ranks tests (Siegel, 1956). Performance of all species on both object and positional discrimination reversals differed significantly ($P < 0·01$) on the initial and final ten reversals. All species showed considerable improvement in performing repeated reversals of both spatial and non-spatial discriminations. Mink attempting the object discrimination reversal showed the most rapid improvement.

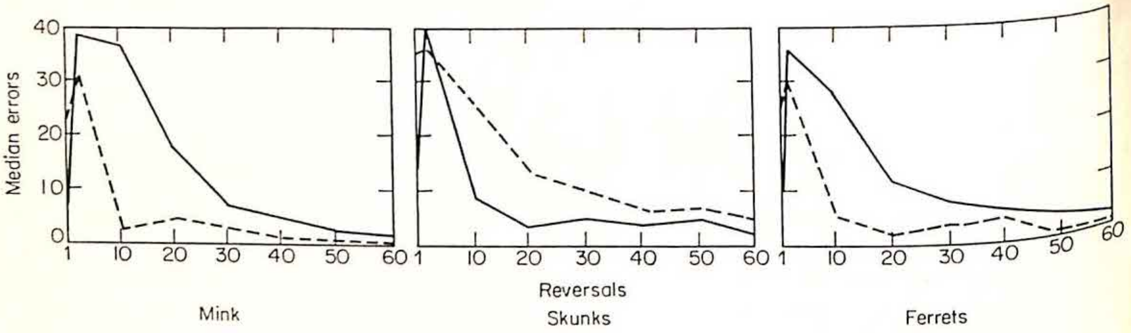


FIGURE 1. Errors to criterion on repeated reversals of an object and positional discrimination by mink, ferrets and skunks. – – –, Object discrimination; ——, positional discrimination.

Total errors to criterion on object and position reversals were analysed by separate Mann-Whitney U Tests for each species (Siegel, 1956). Significant intra-species differences were found in performance of all three species groups on the two types of reversals. Mink and ferrets made significantly ($U = 0$, $P < 0·028$, two-tailed test) fewer errors on object discrimination than on positional discrimination reversals. Skunks learned to reverse the positional discrimination more rapidly ($U = 1$, $P < 0·05$, two-tailed test) than the object discrimination.

The percent of correct responses made by the three species groups on trial 1 to 10 of reversals 1 through 10 and reversals 50 through 60 are presented in Figure 2.

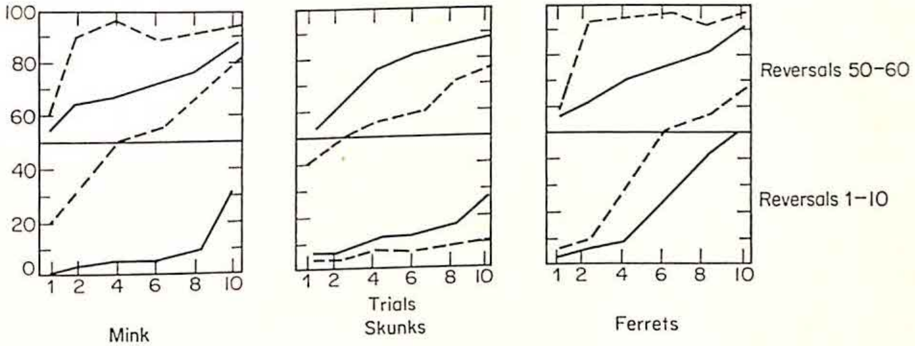


FIGURE 2. Intrareversal learning by mink, ferrets and skunks.

This figure indicates that initial performance on both types of reversals was poor among all species. However, some subjects, namely mink and ferrets, attempting the object discrimination reversal, and skunks trained on the positional discrimination reversal task, exhibited one-trial learning after repeated reversal training. The initial intra-species differences on the two types of reversals persisted to a reduced degree throughout training.

Discussion

It is apparent that certain weasel-family mammals successfully perform repeated reversals of object and positional discriminations with a single pair of stimuli. The discrimination reversal learning curves exhibited by mustelidae species are similar to those noted in a number of other species including primates (Schusterman, 1964; Warren, 1965). These findings suggest that quantitative rather than qualitative differences characterize mammalian discrimination-reversal learning.

Substantial intra-group differences appeared between object and positional discrimination reversal performance. Mink and ferrets approached optimum performance on the object discrimination reversal more rapidly than did skunks. Conversely, skunks learned to perform repeated reversals of the positional reversal more rapidly than did other subjects. Other researchers have found that cats have difficulty learning non-spatial discriminations (Warren, 1965). Skunks, like cats, may be more responsive to position than to object cues. Performance of ferrets and mink on the two types of discriminations suggests that these species, like certain primates, respond more strongly to object than to positional cues (Harlow, 1951; Warren, 1965, 1966). However, evidence summarized by Warren (1965) indicates that species differences in responsiveness to object and positional cues are not closely related to differences in complex learning ability.

This work was supported by NSF URP Grant GE 3943.

References

BACON, H. R., WARREN, J. M. and SCHEIN, M. W. (1962). Non-spatial reversal learning in chickens. *Anim. Behav.* **10**, 239–43.

BITTERMAN, M. E. (1963). Species-differences in learning: Fish, reptile, bird and rat. Paper read at *Am. Psychol. Assn.*, Phila., Pa.

CRONHOLM, J. N., WARREN, J. M. and HARA, K. (1960). Distribution of training and reversal learning by cats. *J. genet. Psychol.* **96**, 105–13.

DUFORT, R. H., GUTTMAN, N. and KIMBLE, G. A. (1954). One-trial discrimination reversal in the white rat. *J. comp. physiol. Psychol.* **47**, 248–9.

HARLOW, H. F. (1949). The formation of learning sets. *Psychol. Rev.* **56**, 51–65.

HARLOW, H. F. (1951). Primate learning. In *Comparative Psychology*, STONE, C. P. (Ed.), 3rd ed., pp. 183–283. New York: Prentice-Hall.

HUGHES, R. N. (1964). Spatial discrimination reversal and overtraining in ferrets. *Percept. mot. Skills* **19**, 817–8.

KIRK, K. L. and BITTERMAN, M. E. (1963). Habit reversal in the turtle. *Q. Jl exp. Psychol.* **15**, 52–7.

ROLLIN, A. R. (1963). Successive object discrimination reversals by squirrels. Paper read at *East. Psychol. Assn.*, New York, N.Y.

SCHUSTERMAN, R. J. (1962). Transfer effects of successive discrimination-reversal training in chimpanzees. *Science, N.Y.* **137**, 422–3.

SCHUSTERMAN, R. J. (1964). Successive discrimination-reversal training and multiple discrimination training in one-trial learning by chimpanzees. *J. comp. physiol. Psych.* **58**, 153–6.

SIEGEL, S. (1956). *Non-parametric Statistics*. New York: McGraw Hill.

WARREN, J. M. and WARREN, H. B. (1962). Reversal learning by horse and raccoon. *J. genet. Psychol.* **100**, 215–20.

WARREN, J. M. (1965). Primate learning in comparative perspective. In SCHRIER, A. M., HARLOW, H. F. and STOLLNITZ, F. (Eds.), *Behaviour of Non-human Primates: Modern Research Trends*, p. 249–81. New York: Academic Press.

WARREN, J. M. (1966). Reversal learning and the formation of learning sets by cats and rhesus monkeys. *J. comp. physiol. Psychol.* **61**, 421–8.

Received 16 *April* 1968

Q. Jl exp. Psychol. (1969) **21**, 63–66

VISUAL COMPLEXITY AND STIMULUS PACING IN CHICKS

J. DUTCH

Victoria University of Wellington, New Zealand

Two groups were tested for spontaneous alternation using massed and spaced trials but alternation rate was not above chance level. Two experiments then examined the use of visual stimuli as pacers for an approach response. The results indicate that subjects responded to complexity and that complexity may be a more effective pacer than novelty.

Introduction

Numerous studies have examined aspects of the relation between stimulus complexity and behaviour in the rat, under a variety of conditions. The common finding of such studies is that the rat will show a preference for the more complex of two stimuli, for example, Dember, Earl and Paradise (1957) and Baker and Franken (1967).

This research suggests that some characteristics of visual stimuli can be used to elicit and pace an approach response in rats. Pace is used in the sense that some aspect of the stimulus has a directive function.

The purpose of the following experiments was to study the effectiveness of both complexity and novelty as pacers for a running response.

Method

Subjects

Subjects were experimentally naive male and female white leghorn chicks. Their age at the beginning of the experiments was 10 to 14 days.

Subjects were maintained in brooders with food and water *ad lib* on a 12 hr. light-dark cycle. The daylight cycle commenced at 8 a.m. and subjects were run in the apparatus commencing at 10 a.m. Thus no subjects came to the apparatus food deprived.

Apparatus

The apparatus consisted of a single straight runway 30 in. long, 6 in. wide, 8 in. high with an 8 in. × 8 in. goal box at each end. Each goal box was equipped with a Perspex guillotine door. Two additional Perspex guillotine doors were used to divide off the middle 6 in. of the runway so that it served as a start box. The subject was placed into the start box through a trap door in the roof of the runway. The walls and ceiling of the runway were lined with acoustic tiles to make it light proof and sound dampened. A 40 W. light was suspended 3 ft. above the centre of the runway so that both goal boxes were illuminated. The floor of the runway and the goal boxes was lined with brown paper, which was changed after each group of subjects had run one block of trials.

Procedure

Subjects were removed singly from the brooder and immediately placed in the start box; the start box doors were raised and the subject could retrace until it entered a goal box. The goal box door was then lowered and the subject left in the goal box for 3 sec. On massed trials the subject was then returned immediately to the start box and the next trial commenced. After each day's trials, or block of trials, the subject was placed in a covered cardboard chick-box lined with shredded paper and left there until all subjects in the group had been run. The order of running was randomized daily for each group. Subjects did not receive any primary reward in the goal box.

Experiment I

The purpose of this experiment was to establish whether there was any tendency for subjects to alternate when all cues (in the sense of Douglas, 1966) were present. Group I ($n = 10$) received four massed trials on one day; Group II ($n = 10$) received four trials with a 24-hr. intertrial interval. Subjects ran to grey goal boxes at each end.

Results and discussion

Group I (massed) alternated on 16 out of 30 trials; Group II (spaced) alternated on 18 out of 30 trials. Group I entered the right-hand goal box 21 times out of 40 trials; Group II entered the right-hand goal box 23 times out of 40 trials. Neither of these differences is statistically significant.

It is obvious that under these conditions chicks do not exhibit spontaneous alternation nor do they show any significant response bias. This supports the study of Hayes and Warren (1963) who failed to obtain spontaneous alteration in chicks using both free and forced choice methods.

Experiment II

Following Experiment I, a further experiment was designed to examine the possibility of pacing the running response in chicks with visual stimuli of differing complexity. The stimulus cards were 8 in. × 8 in. squares of gloss-white cardboard across which 1 in. wide strips of black plastic tape were attached in the following designs: (a) a single central horizontal strip having equal halves of white ground; (b) two horizontal strips giving equal thirds of white ground; (c) two horizontal strips and one central vertical strip giving equal sixths of white ground; (d) two horizontal and two vertical strips giving equal ninths of white ground. The cards were fitted to the end wall of the goal boxes as required.

Following Dember, Earl and Paradise (1957), stimulus complexity is defined as the number of elements constituting each stimulus. In our stimuli we have an increasing number of elements through designs (a) to (d), the assumption being that this corresponds to an increasing complexity of the visual stimulus.

Method

Twelve subjects were given four 24 hr. spaced trials. The position of the stimuli was randomized for each subject, with the restriction that the more complex stimulus appeared equally often in each goal box.

Stimuli were presented in order of increasing complexity so that, for example, on trial 1 the comparison was between a plain white card and stimulus (a), and on trial 4 between stimuli (c) and (d).

Results

Considered overall the more complex stimulus was chosen on 68·75 per cent of the trials.

TABLE I

Per cent subjects choosing more complex stimulus with corresponding P *values*

Trial	%	*P*
1	66·6	0·194
2	83·3	0·019
3	75·0	0·073
4	50·0	0·613

All *P* values are from Siegel (1956) for the Binomial test.

Any interpretation of the results is confused by the fact that on the first trial a choice must be made between two stimuli which are equal in novelty but not equal in complexity. In trials 2, 3 and 4 complexity is confounded with novelty in that the more complex stimulus is also the novel stimulus. Because of this it is not possible to say whether subjects are responding to novelty or complexity. Experiment III was conducted to examine this problem.

Experiment III

Method

Twenty-five subjects were given four 24 hr. spaced trials. Half the subjects always had the more complex of the two stimulus cards in the right goal box, for the other half this was reversed.

Stimuli were presented in order of decreasing complexity so that, for example, on trial 1 the comparison was between a plain white card and stimulus (d), on trial 2 between stimuli (d) and (c), and on trial 4 between stimuli (b) and (a). On trial 1 the stimuli were equal in novelty, but not in complexity, and on trials 2 to 4 the novel stimulus was the less complex.

Results

Considered overall the more complex stimulus was chosen on 64 per cent of the trials.

TABLE II

Per cent subjects choosing more complex stimulus with corresponding P *values*

Trial	%	*P*
1	76·0	0·007
2	68·0	0·054
3	56·0	0·345
4	56·0	0·345

Conclusions

On the basis of these results it is reasonable to conclude that an approach response by chicks can be elicited and paced with a complex stimulus. Although novelty is also an important factor it has less effect than complexity. Further experimentation is required to examine the relative significance of these two factors. For example, the length of time of exposure to the stimuli might be varied systematically, since it could be argued that with an exposure of 3 sec., a complex stimulus may still retain some novelty on later presentations.

References

BAKER, J. G. and FRANKEN, R. E. (1967). Alternation as a function of drive level and visual complexity. *Psychon. Sci.* **8**, 91–2.

DEMBER, W. N., EARL, R. W. and PARADISE, N. (1957). Response by rats to differential stimulus complexity. *J. comp. physiol. Psychol.* **50**, 514–8.

DOUGLAS, R. J. (1966). Cues for spontaneous alternation. *J. comp. physiol. Psychol.* **62**, 171–83.

HAYES, W. N. and WARREN, J. M. (1963). Failure to find spontaneous alternation in chicks. *J. comp. physiol. Psychol.* **56**, 575–7.

SIEGEL, S. (1956). *Non-parametric Statistics.* New York: McGraw-Hill.

Received 26 *August* 1968

Q. Jl exp. Psychol. (1969) **21**, 67–68

THE DEMONSTRATION OF DICHOPTIC FLICKER EFFECTS BY USE OF POLAROID

M. KINSBOURNE† AND P. J. COUGHLIN

Institute of Experimental Psychology, University of Oxford

A simple way of setting up dichoptic flicker with variable phase relationship is described. Sine wave flashes result, and for these, as for square waves, Sherrington's hypothesis of independent testing for intermittence of inputs to each eye is approximately borne out. Observations of the attributes of these two modes of stimulation at sub-fusion frequencies suggest that no such independence holds for computation of luminous flux over time.

The effect on flicker fusion of varying the phase relationship between flickering light sources presented at corresponding loci to each eye was explored by Sherrington (1904). He found critical fusion frequency to be only very slightly less for the anti-phase than for the in-phase condition. Subsequent work has confirmed a decrement of about 8 to 11 per cent due to anti-phase timing (Baker and Bott, 1951; Foley and Stager, 1965; Perrin, 1954; Thomas, 1955). The percept thus fails fully to reflect a major feature of the distribution over time of light incident upon the visual system. In the in-phase condition, in which the binocular input is synchronous, true intermittency of incident light exists. In the anti-phase condition the incidence of light is constant, varying only in its distribution between the retinae. It follows that decision as to intermittency of input is made prior to confluence of information entering though these two channels, with only a minor degree of cross-talk between them. Separate computations on either side are perhaps compared and averaged, or possibly there is alternation between the two eyes, as proposed by Hochberg (1964) for binocular vision in general.

The following simple device permits a classroom demonstration of these effects. A small DC light source of variable intensity is mounted behind a rotating disc of polaroid. The disc is driven by a variable speed shunt wound DC motor, giving a speed range of 120 to 3000 rev./min. The subject wears a modified pair of B.O.C.‡ goggles in which the glass is replaced by matching polaroid discs mounted between thin Perspex. The orientation of one disc to the other is continuously variable between the in-phase and anti-phase condition of the sinusoidal input stimuli.

The stimulus frequency at either eye is twice the motor speed, which gives range of 4 to 100 c/s. To correct for any change in phase angle due to the variation in the facial contours of the subjects, the goggles are braced with 3/16ths steel rod. Provision is made for adjusting to different interocular distance.

---

† Present address: Duke University Medical Center, Durham, North Carolina 27705, U.S.A.

‡ FGI goggles manufactured by British Oxygen Company Ltd.

The flashes differ from those previously used in that they take sine rather than square wave form. Critical flicker fusion frequency is slightly lower for sine than for square wave stimuli (Ives, 1922).

Observations with this apparatus have confirmed Sherrington's independence hypothesis for fusion threshold and a mean decrement of 12·66 per cent in critical flicker fusion frequency for anti-phase as compared in to-phase stimulation was obtained for four subjects. The findings may thus be generalized to the sine wave condition.

At a given sub-fusion repetition rate, in-phase and anti-phase flicker does not look identical. Anti-phase flicker looks brighter, and appears phenomenally faster. Attempts to adjust anti-phase flicker frequency to match these characteristics of in-phase flicker, used as a standard, gives rise to much variability, alerting the observer to the fact that the two conditions appear qualitatively different. While at least at low frequencies in-phase flicker appears punctuated by clear and discrete "black flashes", anti-phase flicker never actually gives the impression of clear interruption of light input, but only of a variability which at low frequencies takes the form of striking gamma movement. It is clear that at all times both eyes contribute to the ultimate percept.

Kahneman and Norman (1964) have given reason to suppose that separate neural systems respond to onset of a stimulus and to its overall intensity. If stimulus onset rate is computed prior to confluence of information from the two eyes and intensity subsequent to such confluence, these observations become intelligible.

References

Baker, C. H. and Bott, F. A. (1951). Studies on visual flicker and fusion. II. Effects of timing of visual stimuli on binocular fusion and flicker. *Can. J. Psychol.* 5, 9–17.

Foley, P. J. and Stager, P. (1965). The phase difference function in binocular flicker. *Can. J. Psychol.* **19**, 47–55.

Hochberg, J. (1964). A theory of the binocular cyclopean field on the possibility of simulated stereopsis. *Percept. Mot. Skills*, **19**, 685.

Ives, H. E. (1922). Critical frequency relations in scotopic vision. *J. opt. Soc. Am.* **6**, 254–65.

Kahneman, D. and Norman, J. (1964). The time-intensity relation in visual perception as a function of observer's task. *J. exp. Psychol.* **68**, 215–20.

Perrin, F. M. (1954). A study in binocular flicker. *J. opt. Soc. Am.* **44**, 60–9.

Sherrington, C. S. (1904). On binocular flicker and the condition of activity of "corresponding" retinal points. *Br. J. Psychol.* **1**, 26–59.

Thomas, G. J. (1955). A comparison of uniocular and binocular critical flicker frequencies: simultaneous and alternate flashes. *Am. J. Psychol.* **68**, 37–53.

Received 30 *October* 1968

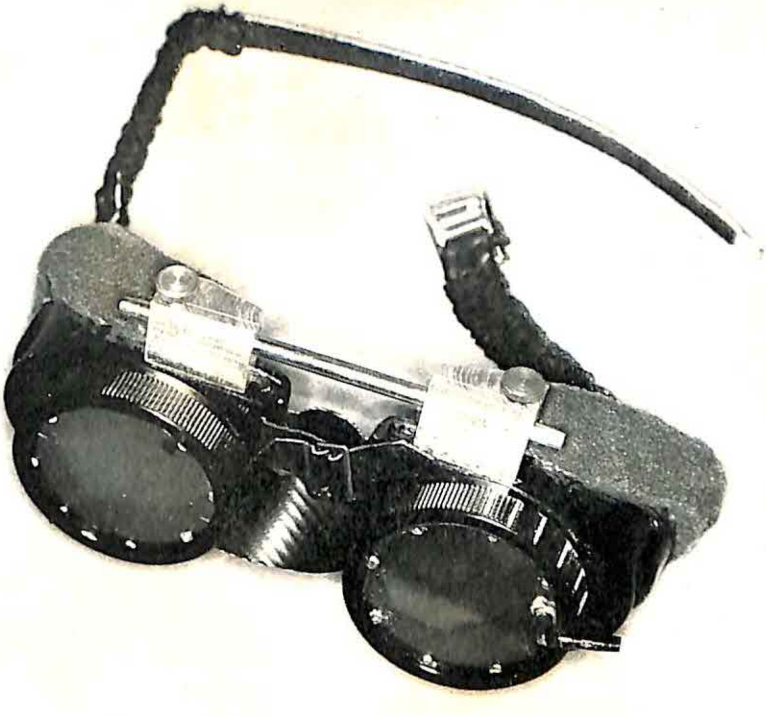


PLATE I. Sketch of modified goggles.

Q. Jl exp. Psychol. (1969) **21**, 69–76

DISCRIMINATION OF "W" AND "V" SHAPES BY GOLDFISH

ROSEMARY BOWMAN AND N. S. SUTHERLAND
Laboratory of Experimental Psychology, University of Sussex

Goldfish were trained to discriminate between "W" and "V" shapes; different groups were trained with the shapes in different orientations. Transfer tests were given after training and the following conclusions were drawn. Animals learned to discriminate between the training shapes by detecting the difference in the number of points present in each; they learned the difference in the *relative* number of points rather than the *absolute* number present in each shape; the subjects transferred well to pairs of shapes bearing points facing in different directions from those on the training shapes; knobs were treated as practically equivalent to points; animals relied more heavily on differences at the tops of the shapes than on differences in the bottom halves.

Introduction

The experiment reported here is one of a series investigating visual shape recognition in the goldfish. There were two main reasons for choosing the training shapes used (depicted in row 1 of Fig. 1). (1) Octopuses find it almost impossible to discriminate between these shapes (Sutherland, 1959). The shapes differ only in their details, and octopuses are very poor at differentiating shapes whose main masses are roughly equally distributed but which have differences in local features. Other evidence suggests that goldfish are quick to discriminate between shapes of this sort (Bowman and Sutherland, 1969) and we tested this further by finding out whether goldfish would be able readily to discriminate between the present training shapes. (2) In the above experiment, it was found that the presence of a knob on a shape could mediate transfer to new rotations of that shape for goldfish, under conditions where such transfer did not occur without the knob. In the present experiment, we tested whether a difference in the number of points present in the training shapes would mediate transfer when further shapes were presented with points facing in different directions from those in the training shapes. It is of some general importance to establish under what conditions transfer to rotated shapes will occur, since very little is known about this.

Method

Subjects and apparatus

Full details and discussion of the apparatus and procedures used can be found in Mackintosh and Sutherland (1963), and in Bowman and Sutherland (1969). Only a brief summary will be given here.

The subjects were 18 common goldfish (*Carassius auratus*), from 3 to 5 in. in length. They were kept in individual tanks, at a temperature of 24°C.

The fish were trained to swim through a hole in a black Perspex detention screen, and take food from one of two troughs attached to the centres of the training shapes. The troughs and shapes were black and were presented on a white Perspex discrimination screen. The two screens were 12 in. apart, and the troughs were 5 in. apart and 2·5 in. from the floor of the tank. The troughs were separated from one another by a vertical Perspex partition projecting 6·5 cm. down the length of the tank. Each fish was trained in its own tank.

Pretraining

All fish were first pretrained to swim through the opening in the detention screen and take a food pellet from a single trough presented on one or other side of the discrimination screen. Each fish was trained until it was regularly taking the food within 10 sec.; this criterion was achieved in an average of 130 pretraining trials. Next, subjects were given 20 trials using the detention screen and a circle (area 17·5 cm.^{2}) behind the trough, to accustom them to the presence of a shape.

Training

Next, training proper was begun, with the positive and negative shapes presented simultaneously. Choice of the positive shape was reinforced with a piece of "Hykro" fish food; the inside of the trough attached to the negative shape contained a blob of paint resembling the food bait. Tests performed at the end of training proved that the fish had learned to discriminate in terms of differences between the two shapes and had not learned merely to discriminate between the food and the blob of paint. The side occupied by the positive shape was determined by Gellermann orders. The criterion of response was the fish's snapping at one of the shapes: this response is unmistakable and easy for the observer to categorize.

At first, if a fish chose wrongly, it was allowed to correct by swimming round to the positive shape. After 20 trials, by which time the fish had lost their initial nervousness of the situation, a modified correction procedure was introduced, in order to check the development of side preferences. When the fish made a wrong choice, the positive shape was shielded by a piece of transparent Perspex inserted between the vertical partition and the edge of the tank on the side of the positive shape. The subject was then re-run until it made a correct choice: if it chose the wrong shape on three successive runs, the negative shape was shielded by a transparent screen, forcing the fish to go to the positive shape. Throughout training, ten trials a day were given, with an intertrial interval of 15 min.

Control and transfer tests

After training was completed, control tests, in which both shapes were baited with food, were alternated with further training trials, in order to confirm that the fish had not learnt to discriminate merely on the basis of any perceptual difference between the food and the paint blob. No correction was given after an incorrect response in a control test.

Next we ran transfer tests alternated with training trials. On each transfer trial a pair of new shapes was shown, both baited with food.

Experimental design

There were three experimental groups, each of six fish; each group was trained to discriminate between a "W" and a "V" shape, but the orientation of the shapes was varied from one group to another. In what follows we shall refer to the training shapes as "W" and "V" shapes but it is important to remember that the shapes were shown in different orientations for different groups. For scale drawings of the shapes used with each group, see row 1 of Figure 1. The shapes were cut from ¼ in. black Perspex, and were 14·4 cm^2. in area. The trough was placed in the centre of each shape. Within each group half the fish were trained with the "W" shape positive, half with the "V" shape positive. The fish were given 200 training trials, including five control tests introduced on the last day of training.

All subjects were given transfer tests with pairs of new shapes. Twelve trials a day were given—six transfer tests and six retraining trials. The shapes used in the transfer tests are also shown in Figure 1. The tests were presented in three series, the first of twelve, the second of six, and the third of two transfer pairs. Within each series, the tests were given in random order; except in the third series, no pair was shown more than once within a day. Ten trials were run with each transfer pair.

Results

Initial learning

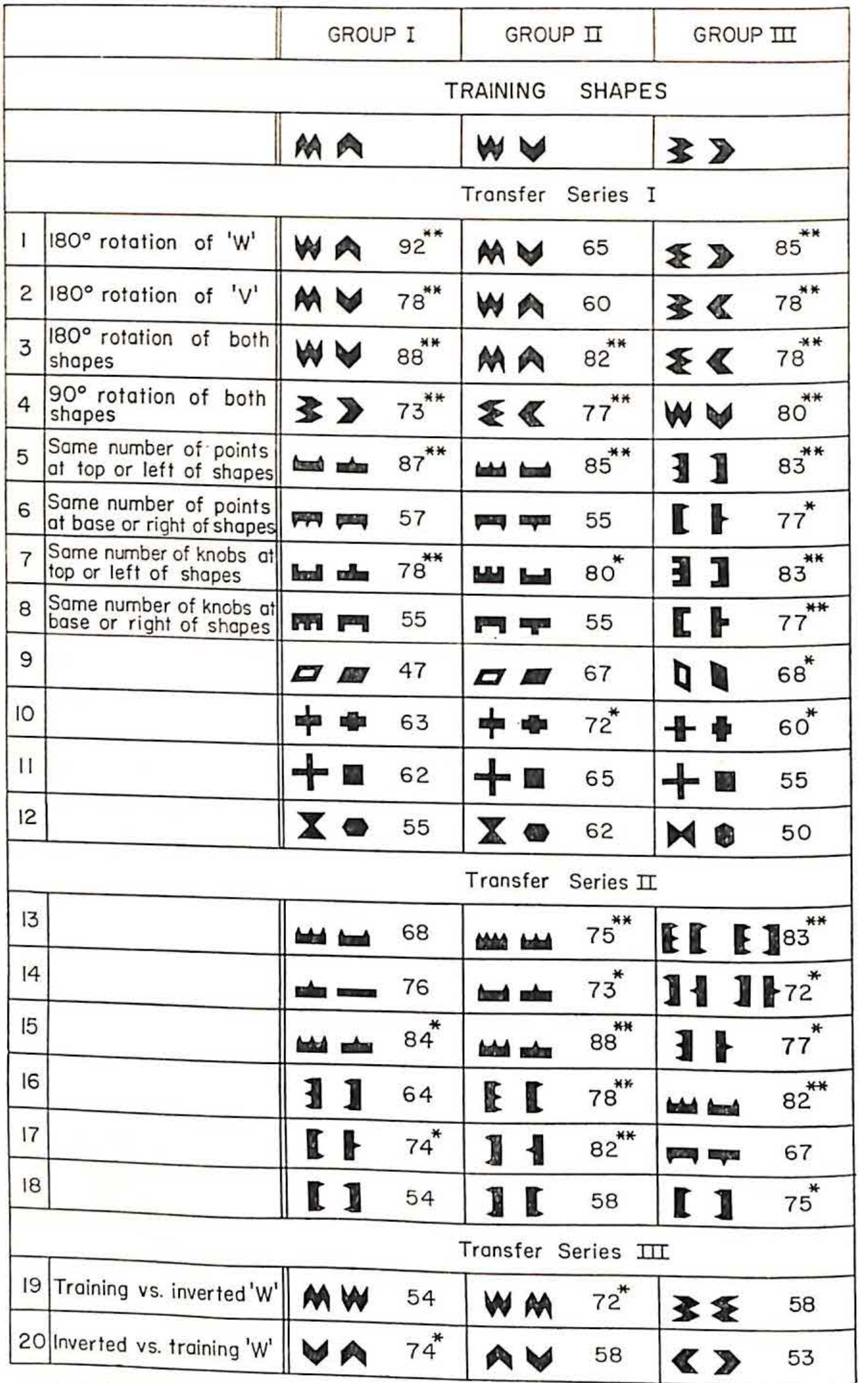
All goldfish met a criterion of at least 18 out of 20 correct responses over a two-day period by the end of training. There were no significant differences among the scores of the three groups, though Group III scored persistently lower than the other two groups, nor was there any difference between the scores of animals trained with "W" positive and those of animals trained with "V" positive. The average numbers of days taken to reach criterion by Groups I, II and III were 13, 11 and 14, respectively. On the twentieth day, the fish averaged 95 per cent correct on the control tests, compared with 93 per cent correct on the intervening retraining trials. The fish therefore, had learned to solve the problem by discriminating differences between the two shapes and had not merely learned to discriminate between food-pellet vs. paint-blob.

Transfer tests

The detailed results of the transfer tests are shown in Figure 1; throughout the retraining trials, the fish maintained a high level of performance (95 per cent correct, see row 1). The numbers in the figure show the percentage of transfer trials on which the appropriate member of the transfer pair was treated as equivalent to the originally-positive training shape. The scores are an average of the percent of trials on which the fish trained with "W" positive chose the left-hand member of the transfer pair (as drawn in Fig. 1) and on which the "V"-positive fish chose the righthand transfer shape. In the second and third transfer series, the Group I scores are based on only five animals, because of the death of one of the "W" positive fish.

The following statistics were calculated on the transfer test data. (1) The probability of all six fish scoring above chance in a given transfer test is 0·016 (0·032 where there were only five fish in the group): in Figure 1, stars are used to indicate these probabilities. Two stars indicate that all six subjects scored above chance, one star indicates that five subjects were above chance and the sixth scored at chance level, or in the case of Group I during Series II and III that all fish were above chance. (2) The significance of the difference between the scores made by the same group (or combination of groups) on two different transfer tests was evaluated by the Sign test. (3) The significance of the difference between the scores of two different groups on the same transfer test was evaluated by the Mann–Whitney U-test. As will be shown in what follows, the results of transfer tests are internally consistent and are therefore valid. Where the same transformation of the training shapes was presented in transfer tests, results for the different groups agree. Where similar transformations of the training shapes were presented on separate tests to the same group, the results also agree.

A cursory glance at the transfer results suggests that animals had learnt the original discrimination by detecting a difference in the number of points present in



		GROUP I	GROUP II	GROUP III
		TRAINING SHAPES		
		Transfer Series I		
1	180° rotation of 'W'	92^{**}	65	85^{**}
2	180° rotation of 'V'	78^{**}	60	78^{**}
3	180° rotation of both shapes	88^{**}	82^{**}	78^{**}
4	90° rotation of both shapes	73^{**}	77^{**}	80^{**}
5	Same number of points at top or left of shapes	87^{**}	85^{**}	83^{**}
6	Same number of points at base or right of shapes	57	55	77^{*}
7	Same number of knobs at top or left of shapes	78^{**}	80^{*}	83^{**}
8	Same number of knobs at base or right of shapes	55	55	77^{**}
9		47	67	68^{*}
10		63	72^{*}	60^{*}
11		62	65	55
12		55	62	50
		Transfer Series II		
13		68	75^{**}	83^{**}
14		76	73^{*}	72^{*}
15		84^{*}	88^{**}	77^{*}
16		64	78^{**}	82^{**}
17		74^{*}	82^{**}	67
18		54	58	75^{*}
		Transfer Series III		
19	Training vs. inverted 'W'	54	72^{*}	58
20	Inverted vs. training 'W'	74^{*}	58	53

FIGURE 1. Training and transfer shapes. The numbers indicate the percentage of trials on which the left-hand member of the pair of shapes shown opposite each number was treated as equivalent to the training "W," and the right-hand shape as equivalent to the training "V."

the two shapes. In describing the results of transfer tests we shall assume that this was true; however, since shapes differing in number of points differ in other aspects as well (e.g. total amount of contour present), more detailed arguments for supposing animals had detected a difference in the number of points present will given be later. The following main conclusions can be drawn from the transfer test results.

(i) *Top vs. bottom halves of shapes.* The animals' performance was determined by differences in the top halves of the shapes much more readily than by differences in their bottom halves. When shapes were presented having the same number of projections in their top halves as the originals, transfer for Groups I and II was very much better than when shapes having the same number of projections in their bottom halves were presented ($P < 0{\cdot}001$ for the significance of difference—compare test 5 with 6, and test 7 with 8). Group III, trained on shapes with points at the side, also did better in transfer when the shapes differed in number of points at the top than when they differed in number of points at the bottom (tests 16 and 17).

However, unlike Groups I and II, Group III did show some transfer ($P = 0{\cdot}062$) to shapes with differing numbers of points at the bottom, possibly because, in order to discriminate the training shapes, Group III was forced to pay attention to the *sides* of the shapes, whereas Groups I and II could concentrate solely on the top halves of the shapes. It is worth noting that early in training, Groups I and II tended to snap at the tops of the shapes, whereas Group III tended to snap at the sides.

(ii) *Position of points.* Given that transfer to shapes differing in number of points at the top was in general better than to shapes with different numbers of points at the bottom, it is necessary to ask whether transfer occurs to pairs of shapes with points facing in different directions from the points on the training shapes. Such transfer clearly did occur, since Groups I and II performed well with shapes differing in number of points at the sides, while Group III, as we have already seen, transferred when the differences occurred at the top or bottom (tests 4, 16, 17).

(iii) *Number of points.* The direction of transfer was determined more by the *relative* difference in the numbers of points than by the absolute number present in each shape. Groups I and II showed good transfer to any pair of shapes having different numbers of points at the top (tests 5, 13, 14, 15); the shape with the greater number of points was always treated as equivalent to the training "W," that with the lesser, to the "V." Transfer was excellent when the absolute numbers of points were the same as at the tops of the two training shapes (test 5); it was also good when the difference in the number of training points was exaggerated (test 15); transfer was significantly worse ($P < 0{\cdot}01$ for Groups I and II combined) when the absolute number of points on each shape was changed but the relative difference in numbers was the same as in the training shapes (tests 13, 14 compared with tests 5 and 15). In tests 13 and 14 we were pitting a difference in the relative number of points against any transfer due to one shape having the same absolute number of points as the original training shape: in test 13, the transfer shape with *fewer* points had the same absolute number of points as the training shape having *more* points at the top, whereas in test 14 the transfer shape with *more* points had the same absolute number of points as were present at the top of the training shape with *fewer* points. Thus, in these two tests, it was the relative difference in number of points that determined transfer, but the fact that transfer was significantly reduced suggests that animals might also have learned something about the absolute number of points present at the tops of the training shapes. When shapes were shown with points at the sides, in positions corresponding to a 90° rotation of both shapes (tests 16, 17), transfer was again good, and was again determined by the relative difference in number of points.

These conclusions are supported by the results of transfer tests 1 to 3 with Groups I and II. Thus in Group I, when one or other training shape was rotated through 180°, transfer remained good provided that the "W" shape had more points at the top than the "V" shape (tests 1 and 3): there was less transfer (though the difference was not significant) when each member of the transfer pair had the same number of points at the top (test 2). Again, for Group II, transfer was poor when both shapes had two points at the top (test 1); transfer was significantly better ($P < 0{\cdot}02$) when one shape had two points, the other one (test 3). However, the results of test 2 with Group II are anomalous in that, when the "W" had three points at the top and the "V" only one, poor transfer occurred: we are inclined to attribute this single anomalous result to chance.

The results of tests 19 and 20 also support the above conclusions. In these tests one or other of the original shapes was presented with its own inversion. When the original "W" had more points at the top than its inversion (Group II, test 19) and when the original "V" had less points at the top than the inverted "V" (Group I, test 20) reasonably good transfer occurred. Performance was no better than chance when the original "W" had fewer points at the top than its inversion (Group I, test 19) and when the original "V" had more (Group II, test 20). However, when inversion of a shape changed the direction of difference in the number of points at the top in this way, we would have expected to find negative transfer, i.e. the inverted version of the training shape should have been treated as equivalent to the original "W" or "V." Since this did not occur, we must conclude that animals had learned something about certain characteristics of the original shapes other than the relative difference between the number of points at the top.

The results do not allow us to decide whether transfer for Group III was determined by the relative number of points present in a transfer pair rather than by the absolute number of points. The reason for this is that, whereas Groups I and II learned much more about the tops of the shapes than about their bottoms, Group III seems to have learned equally about both sides of the shape: thus, there was no significant difference in transfer to shapes having the same numbers of points as were present on the left-hand sides of the training shapes (test 5), and to shapes having the same numbers of points as the right-hand sides of the training shapes (test 6). While Group I learned to discriminate between two points and one, and Group II to discriminate between three points and two, Group III had the opportunity to learn both discriminations from the initial training shapes. The significant transfer obtained in test 18 with Group III suggests that they had learned on which side of each training shape two points appear: this conclusion is weakened by the results of tests 19 and 20, since no significant transfer was shown, although within each pair one member has two points facing right, the other two points facing left.

(iv) *Knobs and points.* Where rectangles with different numbers of knobs were used, transfer was practically as good as to the corresponding rectangles with points (compare test 7 with 5 and test 8 with 6). This rules out the possibility that animals had learned to discriminate between the original shapes in terms of differences in the amount of *oblique* contour present in a particular area. It does not eliminate the possibility that animals were discriminating the difference in total amount of contour or the difference in the ratio $P/\sqrt{A}$, where P is periphery, A is area. This possibility is counterindicated by the results of test 9: although the outline parallelogram contains very much more contour in its top half than the filled-in parallelogram, there was little tendency to transfer from the original "W" to the outline parallelogram. Tests 10 and 11 show that there was some tendency to transfer to a pair of shapes, one of which had a thin segment orientated in the same direction as the points of the original shapes, the other of which had either a thick segment or no segment.

Discussion

One reason for performing the present experiment was to discover how easy it would be for goldfish to master the initial discrimination. Criterion was reached in about 130 trials: the discrimination is therefore more difficult for goldfish than that between a horizontal and vertical rectangle (45 trials—Mackintosh and Sutherland, 1963), and than a circle and square discrimination (85 trials—Bowman and Sutherland, 1969); the present discrimination problem takes almost exactly the same number of trials to master as the discrimination between two opposite oblique rectangles (135 trials—Mackintosh and Sutherland, 1963). The problem is much more readily solved by goldfish than by octopuses, since octopuses fail to discriminate "W" from "V" after 240 training trials, although they learn the horizontal/vertical discrimination in about 50 trials.

Elsewhere (Bowman and Sutherland, 1969) we have shown that goldfish may learn to discriminate between two shapes in terms of the presence vs. the absence of a knob, and we have shown that such a feature may mediate transfer to rotations of the same shape: fish trained to discriminate between a circle and a square showed no transfer to a 45° rotated square, whereas fish trained to discriminate between a circle and a square-with-a-knob showed excellent transfer when the square-with-knob was rotated through 45°. The present results confirm and extend that finding: they show that, where shapes differ in the *number* of points present, transfer will occur to other shapes having different numbers of points when the points are facing in new directions.

A further important finding is that transfer is determined by the relative number of points present in the transfer shapes more readily than by their absolute number of points. Transposition has frequently been found in animals on dimensions such as size and brightness, but as far as we are awared this is the first time it has been clearly shown with differences in shape. The results imply that goldfish can store a highly abstract rule describing an input shape; in the case of Group I, for instance, the "W" shape must have been specified as (amongst other things) "the shape with the more points at the top" rather than as "the shape with two points at the top."

The tendency for animals to rely on differences in the top halves of shapes more heavily than on differences in the bottom halves again confirms earlier results of our own (see Bowman and Sutherland, 1969, for discussion) and also some results obtained by Hemmings (1965) on tropical fish.

This work forms part of a project on Stimulus Analysing Mechanisms supported by the American Office of Naval Research (Contract N 62558-4791). Miss Bowman is supported by a Science Research Council studentship.

References

BOWMAN, R. and SUTHERLAND, N. S. (1969). Discrimination of circles and squares with and without knobs by goldfish. (In preparation.)

HEMMINGS, G. (1965). Pretraining and transfer in shape discrimination. Unpublished Ph.D. thesis: University of Hull.

MACKINTOSH, J. and SUTHERLAND, N. S. (1963). Visual discrimination by goldfish: the orientation of rectangles. *Anim. Behav.* **11**, 135-41.

SUTHERLAND, N. S. (1959). A test of theory of shape discrimination in *Octopus vulgaris* Lamarck. *J. comp. physiol. Psychol.* **52**, 135-41.

Received 22 July 1968

Q. Jl exp. Psychol. (1969) **21**, 77–84

DISCRIMINATION OF CHECKERBOARD PATTERNS BY RATS

N. S. SUTHERLAND AND CLARE WILLIAMS
Laboratory of Experimental Psychology, University of Sussex

Fifteen rats were trained to discriminate between a regular checkerboard pattern and a similar pattern containing a "mistake." The animals subsequently transferred to many other pairs of regular and irregular patterns differing widely from the originals. The results suggest that rats can induce and store highly abstract rules describing input patterns.

Introduction

Sutherland (1968) has recently argued that visual pattern recognition in vertebrates involves two separate processes. (1) The input picture is first decomposed into local features such as edges and corners: the physiological mechanism for detecting such features is known to be present in the vertebrate visual system (Hubel and Wiesel, 1965). The brain preserves the spatial relationships of these local features since there is a topological mapping of the retina onto different levels of feature detectors. (2) The brain stores a description of the feature detectors activated and their spatial relationships. The stored description must be sufficiently abstract to account for the known facts of shape equivalence. That is to say the description must be free from retinal position, and absolute size: the same description must encompass the original input shape, some of its perspective transformations and the mirror-image transform. The rule placed in storage must in fact describe not merely the input shape but a whole class of shapes of which the input shape is a member.

The description forming mechanism almost certainly has built-in biases. In particular, it is likely to take advantage of some types of redundancy in the picture so that a shorter rule will be stored than could be stored by listing independently all the feature detectors fired by a given picture and their spatial relationships to one another. Man can clearly form very abstract rules describing pictures: for example, most adults would tend to classify patterns 1–15 in Figure 1 as similar to one another, and Pattern Reg. as different. The stored description of a regular checkerboard pattern must be highly abstract—we can recognize a checkerboard pattern as a checkerboard regardless of the number of elements, the size of the elements, or whether the pattern starts in the top left-hand corner with a black or white square. Some psychologists (e.g. Glanzer, 1964) have supposed that the human ability to form abstract descriptions of pictures depends upon our possession of a natural language. There are two main advantages in storing an abstract description of a pattern rather than in storing all the details. First, if the abstract description is correctly chosen, it will enable us to recognize an object despite changes in the proximal stimulus. Second, a considerable economy of storage space could be achieved by storing an abstract rule to describe an input picture rather than by listing all the feature detectors fired by that picture. Since these advantages apply as much to other higher animals as to humans, it is possible that they also form highly abstract descriptions of patterns and that this ability in humans is not dependent on language.

The present experiment was undertaken in order to test this possibility. We trained rats to discriminate between a regular checkerboard pattern and a similar pattern containing one irregularity, and then tested them with further patterns to discover to what extent they had learned a general rule describing regular checkerboard patterns. The experiment must be regarded as a pilot study for two reasons. First, as a result of experimental error a mistake was made in choosing the initial patterns to be discriminated. Second, it was difficult to know what transfer patterns to use in advance of gaining some knowledge of rats' performance in this type of discrimination so that after training various arbitrary decisions were taken by the experimenters as to how to proceed next. Though the results are not necessarily completely conclusive, they are thought to be sufficiently new and suggestive to be worth reporting at this stage.

Method

Subjects and apparatus

The subjects were 15 male hooded rats about 100 days old. They were maintained throughout the experiment on a 24 hr. feeding schedule, receiving 1·5 hr. *ad libitum* food after they had been run each day. They were trained on a Lashley jumping stand modified by the inclusion of landing platforms in front of each stimulus window (for fuller description of the apparatus and general procedures see Sutherland and Holgate, 1966). The stimuli used are shown in Figures 1 and 2; they were matt prints measuring 14 × 14 cm. pasted onto the doors of the jumping stand and covering the whole area of the doors.

Procedure

Rats were pretrained to jump to open doors. On training trials, they were rewarded with 10 sec. access to wet mash when they jumped to the positive stimulus. If they jumped to the negative, they were left on the landing platform for 5 sec. and then re-run until they selected the correct stimulus (modified correction procedure). After training, animals were given transfer tests on which at least one of the two patterns presented differed from the training patterns. On transfer trials animals were rewarded whichever stimulus they chose. Transfer trials were always alternated with retraining trials on which the training patterns were presented and animals were only rewarded for correct choices. The initial training shapes were those labelled Reg. and Irreg. in Figure 1. Seven rats (Group Reg.) were trained with Reg. positive (an eighth died), and eight (Group Irreg.) with Irreg. positive. We had intended to use the mirror-image of shape Reg. so that shapes Reg. and Irreg. would have identical elements except at two points where a black and white square are transposed. Our mistake demonstrates the difficulty humans have in discriminating mirror images.

The sequence in which transfer tests and retraining trials were administered will be described in the next section. Where probability levels are given either for differences of animals' scores from chance expectation or for differences between scores on different transfer tests, we applied a sign test unless otherwise stated.

Results

Stage I—Training

All animals were trained with shapes Reg. and Irreg. to a criterion of 18 out of 20 jumps correct over two days of training. Group Reg. met this criterion in an average of 137 trials, and Group Irreg. in 113 trials. The difference is not significant. In this apparatus it takes animals about 70 trials to learn to discriminate between a horizontal and vertical rectangle measuring 10 × 2·5 cm. Since this is known to be a very easy discrimination, the present shapes were moderately easy to discriminate.

Stage II—Initial transfer tests

All animals were given two transfer tests with each of shapes 1–15 (see Fig. 1), paired with the original Reg. pattern: each of these patterns contains a different "mistake," i.e. a different departure from a regular checkerboard. Ten transfer and ten retraining trials were given per day; the transfer patterns were presented in a random order. Group Irreg. performed at chance (50·3 per cent choices of shape Reg.) during transfer tests; Group Reg. chose shape Reg. on 72 per cent of tests—their performance was significantly above chance ($P < 0{\cdot}01$ by a *t*-test). During the retraining trials with the original shapes Group Irreg. were correct on 86 per cent of trials, Group Reg. on 94 per cent. Five of the seven animals in Group Reg. scored 23 out of 30 correct or better on transfer tests: such a score differs from chance with a probability of less than 0·01 by a one-tailed binomial test. These animals scored 97 per cent on patterns 1, 4 and 5 (containing an upright or inverted "T" shape), 78 per cent on patterns 2, 3, 6 and 7 (containing a "T" on its side), and 77 per cent on the remaining figures. If we break up the transfer patterns according to the lowest line in which a displaced element occurred, then animals averaged 97 per cent when there was a "mistake" in the bottom line, and between 78 to 80 per cent for the three other cases ("mistake" on second line from bottom, third and fourth). Although the differences are not quite significant, there is therefore a tendency for transfer to be better when an upright or inverted "T" figure appears in the irregular pattern, and when the mistake occurs towards the bottom of the pattern. Most of the remaining animals developed strong position habits during transfer tests, and Group Irreg. performed poorly even on retraining trials. It was therefore decided to retrain all subjects.

Stage III—Retraining

All animals were retrained using Pattern Reg. paired with Pattern Irreg. and the three 90° rotations of it (Patterns 1, 2 and 3). Pattern Reg. remained as the positive shape for Group Reg., and the four irregular patterns were positive for Group Irreg. Twelve trials were given per day (three with each of the four irregular patterns used) for 28 days. Over the first 24 trials of retraining, animals of Group Reg. averaged 68 per cent correct and of Group Irreg. 54 per cent correct; over the last 24 trials the corresponding figures were 83 and 68 per cent ($P < 0{\cdot}01$ for probability of this improvement being due to chance). Over the whole retraining period animals performed as follows: Shape Irreg.—89 per cent, Pattern

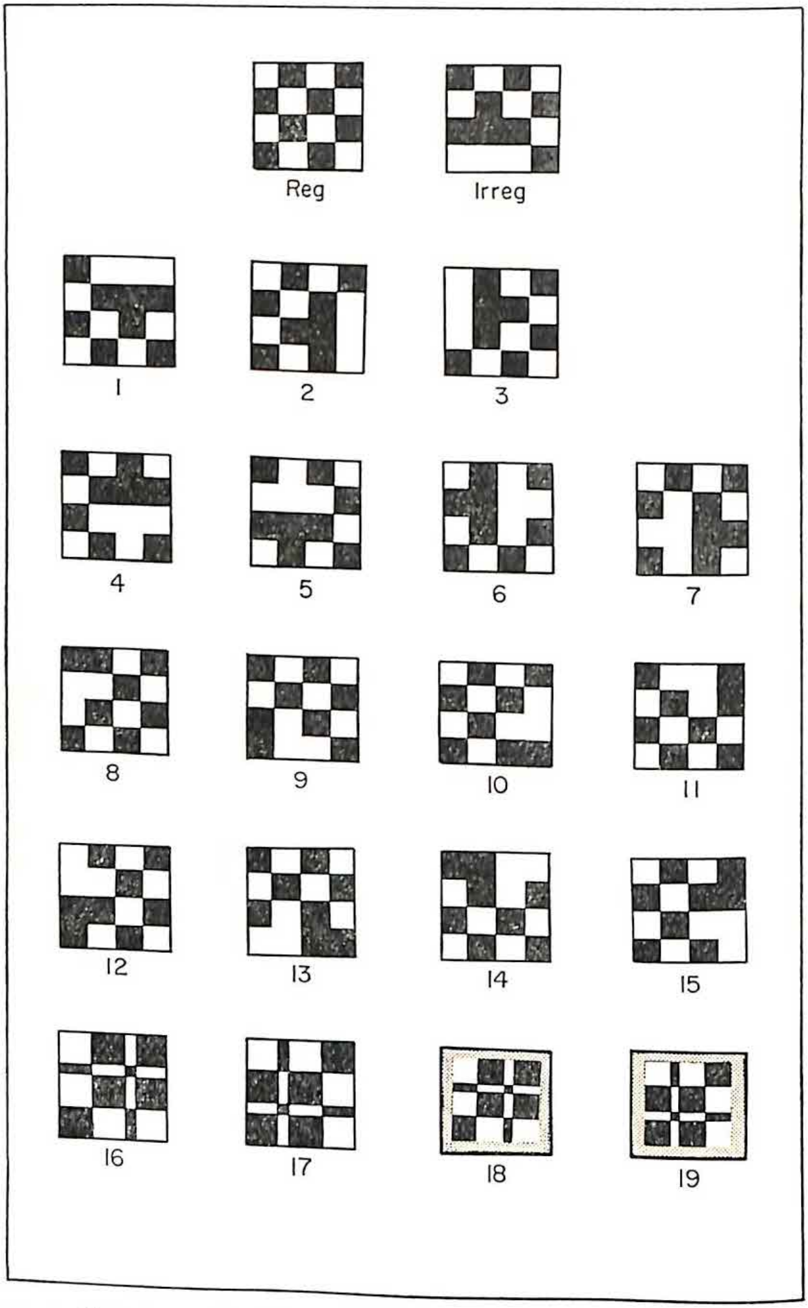


FIGURE 1. Training patterns and initial transfer patterns.

1—73 per cent, Pattern 2—63 per cent, Pattern 3—56 per cent. Performance was significantly better on Pattern Irreg. than on any other ($P < 0{\cdot}001$); it was also significantly better when an upright "T" appeared in the top part of the figure than when a "T" on its side appeared (Patterns 2 and 3, $P < 0{\cdot}001$). These results confirm those from Stage II. Animals performed better when the "mistake" occurred towards the bottom of the irregular pattern and when the irregular pattern contained an upright or inverted "T" rather than a "T" on its side.

Stage IV

At this stage animals were given further transfer tests on patterns 4 to 15 paired with Pattern Reg.: three trials were given to each animal on each pattern interspersed with retraining trials on Patterns Reg. and Irreg. All animals scored above chance during the transfer tests. Group Reg. again performed better on transfer tests than Group Irreg. (84 per cent correct vs. 74 per cent). Once again performance was better when upright and inverted "T" shapes appeared in the irregular patterns (Pattern 4 and 5, performance 89 per cent) than when a "T" on its side appeared (Patterns 6 and 7, 64 per cent, $P < 0{\cdot}002$ for the significance of the difference). Performance on the shapes containing "L"s (shapes 8–15, 80 per cent) was also significantly worse ($P < 0{\cdot}05$) than on shapes 4 and 5. Animals scored 83 per cent correct on patterns where the lowest mistake occurred on the first or second row from the bottom, and 69 per cent correct where the mistake occurred on rows three and four from the bottom—the difference was significant at better than the 0·001 level.

Stage V

In all the transfer pairs described so far, the irregular patterns contain somewhere within them a larger continuous area of black (and white) than do the regular patterns. To discover whether animals were merely discriminating between the two shapes in terms of the presence vs. the absence of such large continuous area of the same shade, further transfer tests were run with irregular Patterns 16 to 19 in Figure 1 each paired with Pattern Reg. These irregular patterns were made up by squashing one row and one column of elements. Since the total size of Patterns 16 and 17 is the same as that of the originals, the size of the square elements is larger: for this reason we included Patterns 18 and 19 which have squares of the same size as those in the training patterns. Twelve tests were run with each pattern.

During transfer tests animals averaged 72 per cent correct (all animals but one were above chance, $P < 0{\cdot}002$). There were no significant differences in amount of transfer to any of the four patterns, the scores on each ranging from 69 per cent to 74 per cent. Group Reg. again performed better than Group Irreg. both during retraining trials (95 per cent vs. 89 per cent) and during transfer tests (81 per cent vs. 64 per cent).

Stage VI

In all the tests reported so far, the original regular pattern has always appeared as one member of the pair. It could be argued that rats had simply stored an exact representation of this pattern and that transfer was being determined entirely by the presence of the original regular figure. To discover whether this was the case a series of further transfer tests was carried out in which both the regular and the irregular patterns differed from the original training patterns.

All animals were given 24 transfer tests with pairs 20 and 21 in Figure 2. One pair is made up of patterns containing 36 elements, the other of 64-element patterns; only 16 elements were present in the original training patterns. A further change was also made to the regular pattern: instead of starting with a white square in the top left hand corner, the regular patterns used started with a black square in this corner. Animals scored 72 per cent ($P < 0{\cdot}003$ for difference from chance performance) on the 36-element patterns, and 58 per cent correct ($P < 0{\cdot}006$) on the 64-element patterns.

Stage VII

Since some rats were developing position habits in transfer tests and hence performing poorly, we decided to retain only the eight animals showing consistently good performance. Three of these eight were from Group Irreg. and five from Group Reg. Twenty-four transfer tests were given to each animal on each of seven new transfer pairs (numbers 22–28 in Fig. 2). On the interspersed retraining trials Group Reg. scored 98 per cent correct, Group Irreg. 96 per cent correct. There were no significant differences between the performance of the two groups on any of the transfer shapes used.

The transfer scores are shown in Figure 2. All animals scored above chance on all tests except tests 23 and 28. Animals did not score significantly above chance on these two tests but on the remaining five tests scores were very significantly above chance ($P < 0{\cdot}005$ on each).

Since all the regular patterns used differed widely from the original Pattern Reg., it is clear that transfer was not being effected by the animals having stored an exact representation of the original regular pattern. Transfer occurred when regular and irregular arrays of triangles and diamonds were presented (Pairs 23–28) and it occurred both when the irregularity introduced a larger continuous area of black and white when it introduced smaller elements. Transfer again tends to be better when the irregularity occurs at the bottom of the shape (compare performance on Patterns 23 and 24, and performance on Patterns 20 and 22). Nevertheless, good transfer can occur when the irregularity is not located at the bottom of the Pattern (25, 26 and 27).

Discussion

The experiment reported was undertaken in order to discover whether rats can learn an abstract rule describing a regular pattern. The transfer test results show that rats do form such rules and suggest that our own ability to discriminate between regular and irregular patterns is not dependent on language. The experiment is in a way a pilot experiment in which we moved from one stage to the next, guided by the results of the previous stage. Some of the results are not as clean cut as we would have liked, but it is felt that the experiment is worth reporting as it stands both because it opens up a new field of enquiry in pattern recognition by animals and because of the significance of the overall results of the transfer tests.

Faults in design

Our mistake in training animals on the mirror image of the regular pattern intended, makes the experiment untidy, but, fortunately we were able to control against the possibility that rats had detected the lack of correspondence between the *regular* elements in the two training shapes: the transfer results clearly show that rats

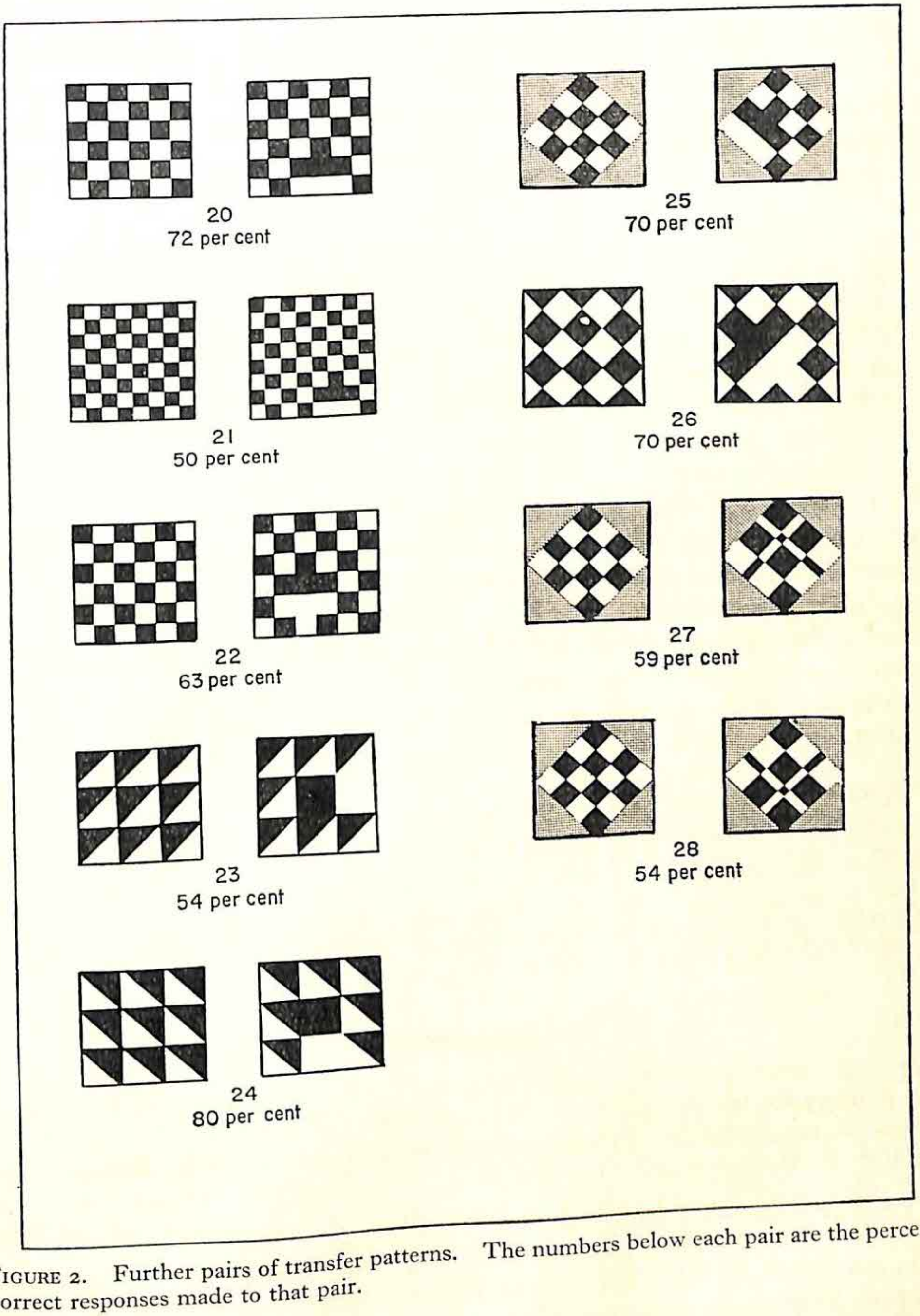


FIGURE 2. Further pairs of transfer patterns. The numbers below each pair are the percentage of correct responses made to that pair.

had not learned the discrimination in terms of the mirror image difference in the upper parts of shapes Reg. and Irreg.

The experiment leaves the following points unresolved.

(1) The subjects may have performed better when irregular shapes contained a mistake in the lower half than when the mistake occurred in the upper half either because the mistake in the training pattern Irreg. occurred in the lower half or

because of rats' well known tendency to detect differences in the lower halves of shapes more readily than in the upper halves.

(2) The subjects performed better when a three element horizontal bar appeared in the transfer shape than when a three-element vertical bar occurred: we do not know whether or not this was because a three element horizontal bar appeared in the original Irreg. training pattern.

(3) We cannot say for certain why Group Reg. showed significant transfer at Stage II while Group Irreg. did not. Since Group Irreg. were trained to jump towards Pattern Irreg. they may have stored more information about the actual locus of the irregularity and hence transferred less well when the locus was changed. Jenkins and Sainsbury (1968) showed that pigeons trained to peck at a pattern containing a feature and not to peck at a pattern lacking the feature, directed their pecks to the positive pattern at the feature. In the same way the rats of Group Irreg. may have orientated themselves to the irregularity.

(4) Finally, we are not in a position to state exactly what rule was formed by rats to describe the regular pattern. The transfer tests eliminate the possibility that the animals had learned to treat any pattern with a large continuous area of black or white as Pattern Irreg. They might, however, have learned that elements of two different sizes appeared on Pattern Irreg. and of only one size on Pattern Reg. though this is counterindicated by the good transfer that occurred with Pattern 26 in which the regular member of the pair contained elements of two different sizes.

Despite these unresolved problems, our experiment does demonstrate unequivocally the capacity of rats to store an abstract description of a pattern: they do not merely store a list of the feature detectors fired by an input picture.

This work forms part of a project on Stimulus Analysing Mechanisms: the project is supported by the American Office of Naval Research (Contract N62558–4791).

References

GLANZER, M. and CLARK, W. H. (1964). The verbal-loop hypothesis: conventional figures. *Am. J. Psychol.* **77**, 621–6.

HUBEL, D. H. and WIESEL, T. N. (1965). Receptive fields and functional architecture in two non-striate visual areas (18 and 19) of the cat. *J. Neurophysiol.* **28**, 229–89.

JENKINS, H. M. and SAINSBURY, R. S. (1968). Discrimination learning with the distinctive feature on positive or negative trials.

MOSTOFSKY, D. I. (Ed.). Stimulus Generalisations. Stanford: Stanford University Press.

SUTHERLAND, N. S. (1968). Shape discrimination in the goldfish. IN INGLE, D. (Ed.) *Central Nervous System and Fish Behaviour*. Chicago University Press, pp. 35–50.

Received 25 July 1968

Q. Jl exp. Psychol. (1969) **21**, 85–93

MEMORY WHILE SHADOWING

DONALD A. NORMAN
University of California, San Diego

Subjects were asked to repeat verbally (shadow) English words which were presented to one ear. They were then tested for their memory of two-digit numbers which were presented to their other ear. It is demonstrated that subjects have no memory for these digits if they must continue to shadow for 20 sec. before being tested on their memory for the digits. However, if tested immediately after the digit presentation, they do remember some digits. Hence, verbal material presented on non-attended channels gets into a short-term memory, but is not transferred into long-term memory.

Introduction

Current theories of attention agree that there is a limitation on our ability to perform several tasks simultaneously: they disagree about the location of the limitation. To simplify the argument somewhat, consider that selective attention operates by disconnecting or attenuating all channels of information arriving at the human except the one of primary interest. The question that arises concerns the level at which this selection takes place: does it occur before or after all channels have been analyzed for their content?

The question arises because signals on non-attended channels can sometimes attract attention if their meaning is important or relevant. Proponents of early selection (Broadbent, 1958; Treisman, 1960, 1964*a*, *b*; Treisman and Geffen, 1967) argue that selection among channels is based entirely on the physical properties of the sensory information and meaning is extracted only from those signals which are selected for further processing. Important signals can usually be detected regardless of the channel on which they occurred because the detection threshold for these signals is sufficiently low to compensate for the attenuation they receive when on non-selected channels. Proponents of late selection (Deutsch and Deutsch, 1963; Norman, 1968) argue that all incoming signals undergo a preliminary analysis of their contents: selection then takes place with the aid of both meaningful and physical properties. The critical distinction between these two classes of theories resides in the way that meaning is extracted from sensory signals. This distinction leads to a testable prediction. The meaning of a physical signal cannot be determined without reference to memory, for the meaning which we attach to the somewhat arbitrary visual and sound patterns of speech and writing must be stored in permanent memory. If meaning is extracted from all signals, then everything presented to us must get analysed through permanent memory and must also reside briefly in short-term memory. This theory must predict, therefore, that we can remember something about all events, even if we fail to pay attention to them. The present experimental literature argues against this possibility.

In general, we remember little of non-attended tasks. This result appears to have been stated first by Cherry (1953) and studied in greatest detail by Moray (1959). In particular, Moray asked subjects to listen to two channels of auditory information presented dichotically while verbally repeating (shadowing) the material heard on one channel. A list of common English words was repeated 35 times over the non-shadowed channel. Within 30 sec. of the end of the shadowing task (which continued for a while after the last presentation of the word list), the subjects were given a recognition memory test for words on the non-attended channel: the false recognition rate was about the same as (and slightly higher than) the correct recognition rate, implying absolutely no memory for the words. A useful control of his experimental procedure comes from the consideration that when the subject's own name was presented on the non-attended channel, it was often recognized.

Several other investigators have examined memory for events on the non-shadowed channel, but in the few cases where memory was found, it was possible to explain this by appealing to a possible or actual disruption of the shadowing task (Mowbray, 1962, 1964). (However, one hint that a brief memory may exist comes from the experiments of Peterson and Kroener, 1964.)

These results, though impressive, do not settle the issue. Under the theories of early selection, failure to remember material on the non-shadowed channel results from the selection occurring during the perceptual analysis of incoming information. To understand the predictions of the theories of late selection, we must first consider what processes are involved in the storage of information and, in particular, the interaction between short- and long-term memory

It is a common assumption that conscious attention is necessary to enact the transfer of material from short- to long-term memory (Waugh and Norman, 1965). If this is so, then we would expect that material on non-attended channels might get into short-term memory but, as long as the subject was occupied with the material on the attended channel, nothing would get into long-term. The data analysed by Waugh and Norman suggest that as few as six to ten items occurring between the presentation of an item and its test reduce the chance that it will later be recalled to zero. Accordingly to this view it would only take a few seconds of shadowing to erase short-term memory completely.

The purpose of this paper is to demonstrate that material presented to the non-shadowed channel is, indeed, stored in short-term memory, but that it never gets to long-term memory. It is not sufficient simply to demonstrate that it is possible to have a memory for items presented on the non-shadowed channel, for that might simply prove that the experiment was poorly done. It is also necessary to prove that there is no memory when conditions are made similar to those used by previous experimenters. Because the results and interpretation of these types of experiments are critically dependent on the experimental procedure, these details will be described in more detail than is usual.

Procedure

Subjects were asked to shadow material presented to one ear, making as few mistakes as possible. Simultaneously, a list of two-digit numbers was presented to their other ear. At a signal, the subjects stopped shadowing and immediately listened to the memory channel, where a test number was presented. The subjects were to decide whether that test number had appeared in the preceding list of numbers. The details of the shadowing task were chosen to make it difficult to perform and the subject's performance was scored. The details of this task and of the scoring procedure were similar to those used by Mowbray (1964). The memory task was chosen to be one about which much is known, and was similar to the recognition experiments of Norman (1966) and Wickelgren and Norman (1966).

One important concern in these experiments is the accuracy of the subjects in shadowing the material presented to them. Some experimenters believe that subjects should produce errorless performance, for errors might indicate a lack of complete attention to the task. This point seems questionable, for if subjects are able to perform without errors, who is to say that their capabilities are being stressed to the fullest? Thus, it would seem wise to design a task that leads to some errors, even when subjects are working at their very best. But, it is important to get good baseline figures on shadowing ability when no other task is required. Then, we can tell if there is any disruption in shadowing performance when the memory task is added to the situation by seeing if there is any disruption of shadowing beyond the baseline established previously. Because of the critical nature of the shadowing performance in this experiment, it seems desirable to use well-trained subjects who have reached a stable level of performance on both memory and shadowing tasks when either is performed alone.

Conditions

Four experimental conditions were studied: shadowing (S), memory (M), memory-while-shadowing (MS), and memory-interference (MI). In condition S, subjects were asked to shadow only and to disregard the memory material. In condition M, subjects were asked to do the memory experiment only and to disregard the words they heard for the shadowing task. In condition MS, subjects were asked to perform both memory and shadowing tasks, but were urged to avoid errors in shadowing. Thus, by instruction and by frequent discussion of the number of errors they made while shadowing, subjects were forced to divide their attention unevenly, with most emphasis on the shadowing task. (The importance of this procedure was discovered during a pilot study in which the subject's performance on the memory task showed little decrement due to the addition of shadowing. His shadowing performance, however, was severely affected by the addition of the memory task. This type of result, similar to that of Mowbray (1964) gives no information relevant to the question asked in this paper.)

Three different sessions of memory material and three sessions of shadowing material were recorded on magnetic tapes making nine possible experimental combinations. During the experiments, the shadowing and memory tapes were played to the subjects over two professional quality tape recorders and earphones. Material to be shadowed was always presented to the left ear and the memory test to the right ear. The subjects sat in individual sound-insulated cubicles. They responded to the memory task by punching their answers onto IBM cards by means of a portable card punch and to the shadowing task by speaking into a microphone. When scheduling permitted, subjects were run at the same time. Each subject participated in a total of approximately 40 experimental sessions (including practice), each session lasted about 1 hr.

Memory task (M, MS, MI)

In the memory experiments, subjects listened to lists consisting of six two-digit numbers. The digits were spoken by a digital computer in a male voice so that the timing and pronunciation of the various conditions was accurately controlled.† Each digit lasted exactly

---

† The digits were made by taking recorded segments of the voice of Douwe Yntema, adjusting them to last exactly 200 msec., and converting the analogue voltage waveform to a digital representation at a sampling rate of 10,000 Hz with 9 bit accuracy. The digits were spoken by putting them out through a digital to analogue converter directly to the tape recorder and thence to earphones. Only the digits zero through seven were used; with the exception that these digits were not compressed in time, the procedure was excatly the same as used in Norman (1966, p. 372). I thank Douwe Yntema for making punched paper tapes of these digits available to me.

200 msec., the two digits of a number were separated by 100 msec., and the numbers were presented at the rate of one digit pair per second. There were two conditions studied: *delayed* and *immediate*. In the *immediate* condition, at the end of the digit list, a 1000 Hz tone (also computer controlled) was presented for 500 msec., followed immediately by a test pair of digits. The subject was asked to judge whether the test number had occurred in the previous list by answering either *yes* (the test number had occurred) or *no* (the test number had not occurred) and then stating his confidence in the accuracy of his response on a four point scale.

In the *delayed* condition, there was a 20-sec. delay between the end of the number list and the presentation of the tone and test number. All other aspects were the same.

Each session of the memory material contained 100 trials, divided into two parts of about 20 min. each. The first 5 trials of each part were used for practice (this fact was not told to the subjects) and so were not counted. Of the remaining 90 trials, 60 were for the *immediate* condition and 30 for the *delayed* condition. Of the 60 immediate tests, 40 were *old* test items (10 at each serial position tested) and the remaining 20 were *new* test items. Of the 30 *delayed* tests, 20 were *old* test items (5 for each serial position tested) and the remaining 10 were *new* test items. To shorten the duration of the experiment only four serial positions in the list were tested: postiions one, three, five and six. The ordering of the conditions and tests was determined by the pseudo-random number generator of the computer: the subject could only tell what condition was being tested at the time of the test itself.

Shadowing task (*S*, *MS*)

Each session of the shadowing material consisted of two parts of about 25 min. each, containing a continuous stream of about 3000 monosyllabic words spoken in a female voice at the rate of two words each second. The words were selected randomly from a list of 2000 words selected from the Thorndike-Lorge collection by Waugh (1961).

Subjects were trained to shadow accurately. They never knew when in a session their shadowing performance was under surveillance, but in fact (because only one subject was monitored at a time) only one-half of any subject's performance was scored. Scoring followed a rigid criterion: any deviation in pronunciation was scored as an error (we followed the rules described by Mowbray, 1964). Occasionally an independent scorer also evaluated the shadowing performance, with very high correlation to the experimenter's scores. Subjects were often shown the results of the shadowing analyses and were continually warned that they were making too many mistakes.

Joint tasks (*MS*)

When performing the two tasks simultaneously, subjects started shadowing the words as soon as they heard them. They were especially warned against increasing their error rate in shadowing when the memory digits were presented. They were instructed to continue shadowing until they heard the 1000 Hz tone in the ear which contained the memory digits. As soon as they heard the tone, they were to stop shadowing, listen to the test digits, and make their response. Immediately after making a response, they were to start shadowing again. There was enough time between lists so that the subjects could do about 5 sec. of shadowing before the start of each list. (The signal to stop shadowing was very effective: missed responses on the memory task were very rare.) This experimental procedure takes advantage of the fact that subjects seem always to detect a tone, regardless of what else they are doing (Cherry, 1953).

Memory-interference task (*MI*)

This task was added in order to study the effect of the shadowing task as pure interference to the memory of the digit lists. Subjects were asked to listen to the digits of the memory task without shadowing. When the six numbers of the digit list had finished, it a tone and test number appeared immediately, they were to respond as well as they could. Otherwise, they were to start shadowing immediately and continue during the 20 sec. that would then elapse between the end of the digit list and the presentation of the tone and test number. In condition MI, therefore, the shadowing acted solely as interference to the memory in the delayed test condition, much like any other interfering task such as colour-naming or counting backwards.

Practice

Subjects practiced shadowing prior to the start of the experiment until their error rate had reached a stable, asymptotic value. This took about 18 hr. of practice, distributed over 9 days. Subjects did several sessions of the memory task prior to the actual experiment to get them used to that condition. In addition, as already noted, the first five memory trials at the start of the day and after each rest were considered to be practice and were not scored.

Subjects

Because of the extreme difficulty and amount of time required to train and test subjects, only two were used in this experiment. (A third subject was used in a pilot study.) They were both female undergraduate students at the University of California, San Diego, who were paid for their services.

Results

The analysis of the shadowing performance is shown in Figure 1. Approximately 10 sec. are required for the error rate while shadowing to rise to an asymptotic value.

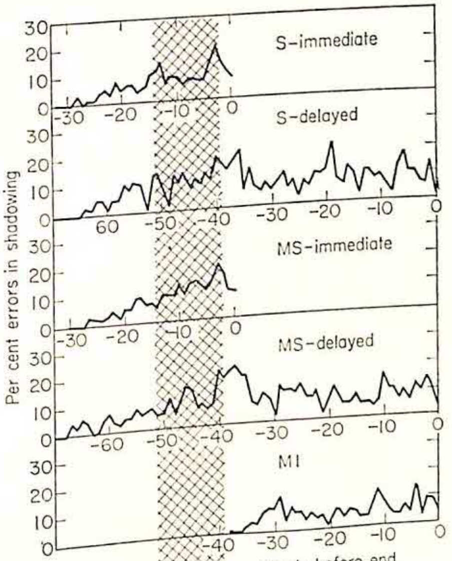


FIGURE 1. The analysis of performance on the shadowing task. The percentage errors are conditional upon shadowing taking place. Thus, the absence of a line indicates that subjects had stopped shadowing.

The shadowing performance, on the whole, is relatively independent of the memory task, although there is a noticeable increase in the error rate at the conclusion of the digit list.

It is not possible to compare the results from the various memory conditions directly because the false recognition rates varied widely. This is a typical result of experiments which use the so-called "yes—no" procedure, whether they be experiments in memory or in psychophysics (see, for example, Green and Swets.

1967; Norman and Wickelgren, 1965; Parks, 1966). For the present analysis, it is convenient to treat the memory task by the assumptions of strength theory (Wickelgren and Norman, 1966). The yes—no and confidence ratings of the subjects give us memory operating characteristics. From these, we can get estimates of memory trace strength for each item. (These estimates of trace strength for some item i, $t(i)$, are similar to the d' of signal detection theory.)

The data for the individual subjects were analysed separately. However, because there was essentially no difference between them, only the average results are reported here. The results of the trace strength analysis are shown in Figure 2.

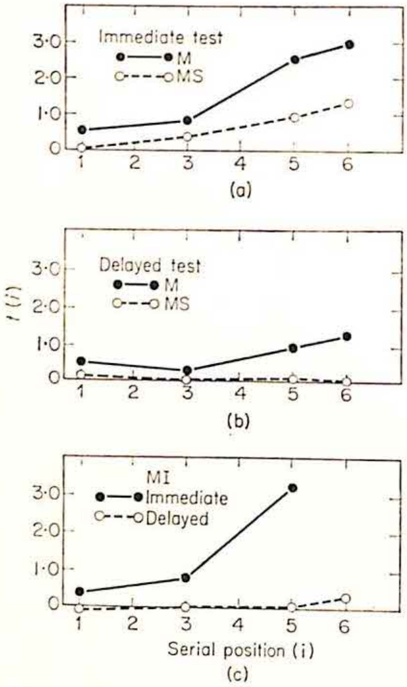


FIGURE 2. The comparison of memory trace strength, $t(i)$, with and without shadowing, conditions MS and M, respectively.

(a) Shows substantial short-term memory in both conditions. (b) Shows substantial long-term memory when the test is delayed 20 sec. after the end of the list but no shadowing is required (condition M, ———). When there is shadowing (condition MS), there is a negligible amount of long-term memory trace strength. (c) Shows that when shadowing takes place only in the interval between the end of the memory list and the test short-term memory (———) is the same as condition M of part A but long-term memory trace strength is completely negligible, as it is in condition MS of part B. (Trace strength for serial position 6 of the immediate test is too high to be estimated.)

The probabilities of correct recognitions, P(y| old), and false recognitions, P(y| new), are shown in Table 1.

Consider first the *immediate* tests of memory. Each estimate of correct recognition, P(y old), represents from 110 to 120 observations and each estimate of false recognition, P (y new), from 200 to 240 observations. It is clear that even though the *immediate* memory is not as good while shadowing (condition MS) as in the absence of shadowing (condition M)—memory strength for position 6 drops from 3·0 to 1·4—there is no question that subjects remember a substantial amount of material from both conditions (Fig. 2(a)). It is also clear that similar serial position effects are observed in both cases. Now consider the *delayed* tests of memory. Each estimate of correct recognition represents from 50 to 60 observations and each estimate of false recognition from 112 to 120 observations. Although there is a substantial amount of memory and a serial position effect when memory is tested in the absence of shadowing (M), there is very little memory at all for the condition with shadowing (MS) (Fig. 2(b)). In condition MS, with *delayed* test, the correct recognition rate is just barely superior to the false recognition rate: the average trace strength of memory items is only 0·11—a very low value indeed. The conclusion one reaches, therefore, is that while shadowing there is a fair amount of *immediate* memory but that almost no memory remains after a delay of 20 sec.

TABLE I

Recognition probabilities in the memory tasks

Condition	Serial position				
	1	3	5	6	new
M-immediate	41	50	92	98	22
M-delayed	61	52	71	86	35
MS-immediate	48	59	83	83	48
MS-delayed	70	68	73	66	62
MI-immediate	43	58	98	100	24
MI-delayed	60	63	68	75	69

The short-term memory measured while shadowing (condition MS) is at a lower level than that observed in the absence of shadowing (condition M). There are at least two reasons for this. First, when the subject is shadowing, he is so busy that he cannot rehearse the digits, even though they are registered in short-term memory. Second, the subjects do not stop shadowing the instant the tone for the test digit is presented. As can be seen in Figure 1, subjects continued to speak for several words after the conclusion of the digit list. These two factors—the interference of the shadowing after the end of the digit list and the lack of rehearsal—both tend to lower the amount of material that is available in memory at the time of test, even if the digits themselves get in equally well whether or not shadowing is taking place. Indeed, the shadowing task is so complex, both in the skills required of the subject and the confusion of sounds that result, that the remarkable point is that there appears to be so little decrement of performance in the memory task. Peterson and Kroener (1964), for example, suggest that the confusion of voices may create a perceptual deficit.

Proof that the lack of memory in the MS-*delayed* condition results from the act of shadowing during the 20-sec. delay comes from a consideration of condition MI, where shadowing was performed only during the delay period (Fig. 2(c)). In the immediate test condition, each estimate of correct recognition represents 80 observations and each estimate of false recognition 160 observations; in the *delayed* test condition, each estimate of correct recognition represents 40 observations and each estimate of false recognition 80 observations. The results of MI-immediate are very similar to those of M-immediate condition, as well they should be, for the conditions are identical. The results with *delayed* tests show slightly less memory for condition MI for condition MS (average memory strength is 0·07 compared with an average of 0·11 for condition MS) but a slightly greater serial position effect. These memory effects are minute, however. For example, the average probability of responding yes to an *old* item is 0·66: this is slightly *less* than the false recognition probability of 0·69. (The overall positive value for memory trace strength comes about through a consideration of confidence ratings.)

Discussion

While a subject is busy shadowing, he still retains a good short-term memory for other material which may be presented to him, but no long-term memory for that material. The experimental conditions used here which are most similar to those of previous studies is the *delayed* test of memory: the results are also very similar—little or no memory. The difficulty with the earlier studies appears to be in their neglect of short-term effects, which are now revealed to be rather large. This neglect is not surprising, for the earlier studies occurred before the recent extensive theoretical and experimental analyses of short-term memory. The present results must be interpreted to indicate that, although the act of shadowing prevents the subject from performing whatever operations (e.g. rehearsal) are necessary to retain material for reasonable durations of time, the short-term memory is still operative.

What do these results imply for theories of attention that assume that we turn off or attenuate all but a single channel of information? Accordingly to these theories, if full attention is paid to the shadowing task, the memory task of the present experiment must be "disconnected" or attenuated. The datum that must be explained by all theories is the short-term memory in the presence of divided attention.

As we have seen, one theory requires the existence of short-term memory. One way for other theories to handle this result is to postulate that there are sensory storage units prior to the attention mechanism. After the signal to stop shadowing has been presented, subjects are able to read out the contents of that sensory storage. When there is a long delay between the end of the memory list and the test, the material in the sensory store decays away. Such a sensory storage system plays an important role in Broadbent's early theories (1958). The main problem with this argument comes from the similarities of the properties of the sensory storage which would be required to handle the present results with the properties of the usual short-term memory. The characteristics of short-term memory usually seem to depend upon such non-sensory factors as the number of psychological units or "chunks" presented rather than upon such sensory factors as the duration of the material (Miller, 1956; Norman, 1966; Waugh and Norman, 1965). The quantitative similarity of the serial position curves from the short-term memory conditions of this study to those found in other studies (such as Norman, 1966 and Wickelgren and Norman, 1966) would appear to be too much of a coincidence to be explained by the parameters of a sensory memory. Further discussion of this point can be found in Norman (1968, 1969).

In conclusion, the experiments reported in this paper illustrate that even while subjects are kept busy with a difficult shadowing task, they retain some short-term memory for other items presented to them. Little or no memory for these items remain after a lapse of 20 sec. The lack of long-term effects can be attributed entirely to the act of shadowing during the 20-sec. delay quite independent of the shadowing that took place at other times. These results are of critical importance for theories that postulate that all sensory signals undergo processing; a negative result would have ruled against these theories.

Supported by research grant NB07454 from the National Institute of Neurological Diseases and Blindness, U.S. Public Health Service. I thank Peter Lindsay for his assistance with the manuscript and Leanne Hinton for her aid in performing and analysing the results of the experiment.

References

Broadbent, D. E. (1958). *Perception and Communication.* New York: Pergamon Press.

Cherry, E. C. (1953). Some experiments on the recognition of speech, with one and with two ears. *J. acoust. Soc. Am.* **24**, 975–79.

Deutsch, J. A. and Deutsch, D. (1963). Attention: some theoretical considerations. *Psychol. Rev.* **70**, 80–90.

Green, D. M. and Swets, J. (1967). *Signal Detection Theory and Psychophysics.* New York: Wiley.

Miller, G. A. (1956). The magical number seven, plus or minus two. *Psychol. Rev.* **63**, 81–97.

Moray, N. (1959). Attention in dichotic listening: affective cues and the influence of instructions. *Q. Jl exp. Psychol.* **11**, 56–60.

Mowbray, G. H. (1962). Some human perceptual limits. *Johns Hopkins Univ. Appl. Phys. Lab. Tech. Dig.* **1**, 21–4.

Mowbray, G. H. (1964). Perception and retention of verbal information presented during auditory shadowing. *J. acoust. Soc. Am.* **36**, 1459–64.

Norman, D. A. (1966). Acquisition and retention in short-term memory. *J. Exper. Psychol.*, **72**, 369–81.

Norman, D. A. (1968). Toward a theory of memory and attention. *Psychol. Rev.*, **75**, 522–36.

Norman, D. A. (1969) Memory and Attention. New York: John Wiley.

Norman, D. A. and Wickelgren, W. A. (1965). Short-term recognition memory for single digits and pairs of digits. *J. exp. Psychol.* **70**, 479–89.

Parks, T. E. (1966). Signal detectability theory of recognition memory performance. *Psychol. Rev.* **73**, 44–58.

Peterson, L. R. and Kroener, S. (1964). Dichotic stimulation and retention. *J. exp. Psychol.* **68**, 125–30.

Treisman, A. M. (1960). Contextual cues in selective listening. *Q. Jl exp. Psychol.* **12**, 242–8.

Treisman, A. M. (1964a). Verbal cues, language, and meaning in selective attention. *Am. J. Psychol.* **77**, 206–19.

Treisman, A. M. (1964b). The effect of irrelevant material on the efficiency of selective listening. *Am. J. Psychol.* **77**, 533–46.

Treisman, A. M. and Geffen, G. (1967). Selective attention: perception or response? *Q. Jl exp. Psychol.* **19**, 1–16.

Waugh, N. C. (1961). Free versus serial recall. *J. exp. Psychol.* **62**, 496–502.

Waugh, N. C. and Norman, D. A. (1965). Primary memory. *Psychol. Rev.* **72**, 89–104.

Wickelgren, W. A. and Norman, D. A. (1966). Strength models and serial position in short-term recognition memory. *J. math. Psychol.* **3**, 316–47.

Received 30 *September* 1968

Q. Jl exp. Psychol. (1969) **21**, 94–103

BOOK REVIEWS

Brain Function. Vol. IV: *Brain Function and Learning.* Edited by D. B. Lindsley and A. A. Lumsdaine. Berkeley: University of California Press (London: Cambridge University Press). 1968. Pp. xiii + 364. 142s. 6d.

This is one of the conference series which records the discussion during and after a paper. I believe that the participants do some editing after the conference, which presumably makes for less liveliness and the two-year publication delay but also for less wasted print. Occasionally the earlier multidisciplinary conferences in this series managed an impressive few pages of truly *inter*disciplinary criticism and integration. Here the cross-disciplinary illumination glows rather than sparkles.

So much of the book's value must be sought in its achievement as a wide but coherent survey of undisciplinary work on the brain in learning. Its success is moderate. The papers are uneven in quality and somewhat disconnected, but the discussions rectify this to some extent. And, anyway, is that not the present state of research into the physiological mechanisms of learning?—different techniques being exploited in ways little related to each other and with varying relevance to the heart of the problem.

The neurochemistry and the electrophysiology of learning have a chapter each, thin on data and critical discussion of it. One could offer a newcomer to the biochemistry of memory a good brief introduction to the present problems, if he cared to work backwards through the discussion of Gaito. Galambos cites without exposition the attempts to find electrical correlates specific to learning by Adey and by John and does not mention the European results. An appendix to the following chapter describes R. F. Thompson and Spencer's analysis of the synaptic basis of spinal habituation in the cat.

There is only one other chapter on physical mechanisms mediating learning. Pribram gives a lucid, interesting introduction to the analysis of the behavioural effects of inferotemporal lesions up to 1964 and his ideas on the role of the "association" areas in corticofugal control of central brain stem sensory input, mediating selective attention and intention.

A number of the other chapters, which confine their data to brain function at the performance level, are also brief and clear introductions to important work. Konorski (with N. E. Miller and Pribram) expounds the experimental dissociation of classical conditioning and instrumental learning and a corresponding distinction in the food system between consummatory and search organizations. Lipsitt describes habituation and some form of conditioning seen in reasonably humane, but well-controlled, experiments with newborn human infants. The psychological, but *not* necessarily physiological, defect(s) in mental retardation are discussed by O'Connor with Zigler and Birch.

There is also Simon on computer programs as fully explicit theories of learning, Atkinson on algebraic models of verbal learning and Lumsdaine on teaching machine technology. Hilgard provides a liberal introduction to positions in qualitative learning theory. Throughout the book, incidentally, S-R theory is treated as a very live horse still enthusiastically backed by others, even though the participants of this conference place their bets elsewhere. Finally, there is Guildford on creativity and Koestler on the act of creation. It is well worth emphasizing that intellectual creation is the summit of human learning, but these authors do not succeed in being as representative of their area as do some other contributors.

As a collection of some good introductions to certain aspects of the analysis of brain functions in learning, which are not obtainable in quite this form elsewhere, the book should probably be in a medium-sized institutional library that included physiological psychology in its range. I cannot imagine an individual finding over £7 worth of usefulness in it on his personal bookshelf, but it could be worth an hour or two picking out the bits that interest (the Name Index includes discussion comments).

D. A. BOOTH

Lehrbuch der experimentellen Psychologie. Edited by R. Meili and H. Rohracher. 2nd edition, revised and enlarged. Bern and Stuttgart: Verlag Hans Huber. 1968. Pp. XII + 501. D.M.44.

It is a long time since a major text in experimental psychology emerged from its birthplace, Germany. This one, first published in 1963 and now reissued in a revised and enlarged edition, attempts to cover the gneeral field of human experimental psychology. The topics considered are psychophysics (Ekman), perception (I. Kohler), learning and memory (Rohracher), thinking (Meili), feeling and emotion (Karsten), personality (Brengelmann), experimental social psychology (Eyferth) and statistical methods (Pawlik). Oddly, perhaps, there is no chapter on language.

In spite of its many excellent qualities, this textbook is pervaded by a curiously old-fashioned sense of the "generalized normal adult human mind" which for Wundt and others of his time represented the subject-matter of experimental psychology. There will be found almost nothing about children or animals, about individual differences in ability or about brains and behaviour. Indeed Behaviourism might almost not have been and Gestalt psychology still maintains a ghostly presence. Clearly, little has changed in German academic psychology except its quality, which seems alas never to have recovered from the *debacle* of 1933.

None the less, one must welcome the Editor's initiative in compiling this text, which does make a good deal of modern work available to German-speaking students. It will no doubt be widely used and in spite of its limitations should keep the tattered banner of what was once the "new psychology" flying in the land of its birth.

O. L. Zangwill

The Measurement and Prediction of Judgement and Choice. By R. Darrell Bock and Lyle V. Jones. San Francisco: Holden-Day. 1968. Pp. 370. $13.75.

Forty years ago, L. L. Thurstone developed a system of psychological measurement based upon the notion that associated with each stimulus was a discriminal process which was a normally distributed random variable. Its dispersion was used to define the unit of a psychological scale. By assumption a person judges A greater than B, or chooses A in preference to B, if the discriminal value of A exceeds that of B. This minimal theory of the processes of judgement and choice led to models for sensory discrimination, comparative judgement and categorical judgement.

Bock and Jones dedicate their book to the master and give a systematic presentation of his models organized under the different methods for obtaining data—the constant method, method of pair comparisons, method of ranks and method of successive categories. Their major contribution, and it is an impressive one, is that they have brought together a wide variety of statistical techniques for solving the problems of parameter estimation, constructing confidence bounds and testing goodness of fit of each model examined. There can be few mathematical models in psychology for which so much attention has been paid to these problems.

For the basic scaling models and for a number of variants they compare graphical, minimum transform chi-square (Urban's method of weighted least squares) and maximum likelihood solutions to parameter estimation and also the confidence bounds on scale values that can be obtained in each case. They also construct appropriate analysis of variance designs for testing goodness of fit of each model.

This book should be useful to anyone concerned with extracting information from data, whether or not his interest is in scaling. He would be advised to first master some text in statistical theory at the level of, say, Mood and Graybill, although he is assisted by worked examples wherever new techniques are introduced. There are nearly 60 pages of tables and computing notes.

If the book is to be faulted we would point to the dissonance between the level of rigour established and the amount of computation this involves on the one hand and on the other hand, the importance of the psychological data discussed. The following detail is typical:

". . . Two samples of warmed green beans were served to customers, who tasted the samples and expressed their preference for one of them. . . ." (p. 141).

HENRY SHAFFER

Minnesota Symposia on Child Psychology. Vol. 1. Edited by John P. Hill. London: Oxford University Press, Minnesota University Press. 1967. Pp. 239. 47s. 6d.

In the last decade, most American work done in the name of child psychology involved taking ideas or techniques from mainstream psychology and trying them out on children; often it seemed, for no better reason than that children are there. Though some of the limits of generality of the adult and animal data became clearer, this era did not add much to our knowledge of development. Now there are signs of increasing concern for what must be the root of child psychology; the problems of development itself. Several influences have contributed to this change. One is certainly the growing knowledge of Piaget's work in the United States. The nativistic approach of the psycholinguists and their growing preoccupation with developmental questions has had its effects, too. Another strong push has come through the boom in studies of the neonatal period. Such research raises many developmental questions in their most acute form. From this perspective the volume under review is antedeluvian in outlook. With the exception of one chapter (by White), one may search in vain for any discussion of developmental problems.

The book has its origins in a symposium held by the Institute of Child Development of the University of Minnesota in 1966. Six speakers were "given the opportunity to discuss a series of related studies from the perspective of the theory or set of issues that guides their programs of research." This, we are told, has unique advantages for both the investigator and the reader, as it provides a "programmatic" perspective which is lost in any "second-person review" of an area and is fragmented in journal reports. The advantages seem a little less than unique as much of the work has been published elsewhere in very similar forms and for the most part the authors have used the opportunity to string together a number of papers with a little bland discussion, apparently unhindered by editing.

In the one example of work from the new wave, White discusses the sensorimotor development of institution children in the first year of life. He points out that we lack the basic observational data necessary to build or test theories of development. Though I feel he underestimates the amount of existing information, I would not quarrel with his plea for us "to sharpen our observational tools and to collect 20 years of solid natural history." He ascribes the dearth of observational work to the difficulties in getting hold of subjects, the lack of research time available for academics and the biases of journals against descriptive papers. These are superficial reasons. Children are easier to come by than say, adult bus drivers, and many animal workers, even in America, find time to do observational work. As for the journals, they merely reflect the work actually being done. The real reasons are slightly deeper than this and may be found in White's chapter itself. Having stressed the need for a biological and observational approach, he goes on to say that he sees the domain of the psychologist as the manipulation of experience through environmental modifications" and he "considers himself professionally obliged to work on the problem of how to structure a child's world so that experience has a maximum positive effect on growth and development." You can't have it both ways. In the long term these may be desirable ambitions, in the short term they are not going to get the observational work done. Natural history is not compatable with structuring the child's world.

This collaborative work (with Hayes, Held and Castle) shows that some environmental modifications influence the development of visual attention and grasping. Giving infants something interesting to look at results in earlier grasping, apparently because they spend less time looking at their own hands. The demonstration of plasticity in these systems is important in itself, but this work goes further as it has opened up a new field for experimental investigation.

Gewirtz, prominant on the S-R right wing of child psychology, demonstrates that the reinforcing properties of the word *good* (or *yafeh* for Hebrew children) in a discrimination learning situation depends inversely on prior presentation.

The rather successful use of operant techniques with schizophrenic children is described by Lovaas. In most treatment situations children are rewarded for psychotic behaviour; the child who bangs his head gets most attention. Lovaas punished psychotic behaviour and used a variety of techniques to shape better things. His claims are modest and he quickly gets you on his side by saying that there is more to human behaviour than that contained in reinforcement theory. The most interesting work concerns verbal imitation training. He began by reinforcing all vocalizations and visual fixations of the adult's mouth and then shaped the behaviour to approximate the model more closely. After several weeks of this, many new words were being learnt each day. At this stage he went on to teach a labelling vocabulary and finally to the training of abstract speech. The degree of generalization of the training is surprisingly wide and seems to stand up reasonably well in other situations. Part of this success may be attributed to the use of the parents in the training sessions. His dismissal of Miller's doubts that operant techniques can produce flexible language seem a little naïve. Precise descriptions of the children's verbal behaviour before treatment are not given. Some were echolalics who must have had some familiarity with speaking before training, in any case more than complete mutes.

The increasing influence of Basil Bernstein's ideas in the United States is seen in the study by Hess and Shipman. I.Q. tests and analysis of tasks carried out by mothers with their four-year-old children form the basic data. The families were all Negro and were divided into four social class groups (college educated, skilled manual, manual and manual with father absent). The three lower class groups differed little among themselves but were distinct from the middle class group in terms of the children's performance and its control by the mother. Though this parallels the mothers' I.Q.s, the authors prefer to discuss the results in terms of social class and the lack of choices for ghetto children and their impulsiveness (delay of gratification means no gratification). This implies that "cultural deprivation" is the lot of any black (and white?) child whose parents don't make it to college.

Mavis Hetherington looks at three theories of sex typing. Her results "appear to yield evidence in partial support of the three theories" (and others too, perhaps?). After four solid pages of intercorrelations of parental and child variables, she remarks that the reader may be dazzled by the array of insignificant correlations. This reviewer was.

Suppes, Hyman and Jerman contribute a chapter entirled "Linear structural models for response and latency performance in arithmatic on computor-controlled terminals." The reviewer does not feel competent to comment on this interesting-looking research except to say that it is refreshing to see the learning theorists getting into the classroom.

The editor remarks that volume illustrates that there are many ways of doing good work in child psychology. But then some people may feel that there is more to the game than *doing* psychology. Though this book is not worth buying, if you are interested in where the child psychologists were a few years ago, it might be worth borrowing from the library for a few days.

MARTIN RICHARDS

An Introduction to Animal Behaviour. By Aubrey Manning. London: Arnold. 1967. Pp. viii + 208. 32s. cloth. 21s. paper.

Part of the Contemporary Biology Series, this is an excellent introduction to animal behaviour. Unlike most introductory texts, it will not mislead the beginner into any narrow path, for animal behaviour is considered from a broad approach, involving ethology, psychology and physiology.

The first chapter in this analysis of the organization of behaviour within the individual is *Reflexes and complex behaviour.* In addition to their differences, it stresses their common properties and provides a brief physiological background for what follows. How an animal becomes well-suited to its environment is considered in the *Development of behaviour.* Examples of selective learning in particular species will be of especial interest to psychologists. Although this chapter gives "little encouragement for attempts to classify behaviour into that which is inherited and that which is learnt," the terms are nevertheless used as if a dichotomy existed, with the danger of the beginner accepting the latter. Factors which determine what an animal does at any given moment include *External stimuli* (sensory capacities, sign stimuli, stimulus filtering) and *Motivation* (from theories of motivation to the physiological basis of motivation). *Conflict behaviour* involves what happens when the smooth course of behaviour is interrupted (threat dispalys, displacement activities, courtship and prolonged conflicts). With *Hormones and behaviour*, a brief physiological basis is again provided before the interactions between hormones and behaviour are considered. The last part of the chapter on insect reproductive behaviour provides a valuable complement to the previous discussion: "it shows that animals with a completely different structure and evolution from the vertebrates nevertheless use similar methods to co-ordinate their behaviour with their reproductive physiology." The following chapter, *Evolution*, concerns genes and behaviour, the micro-evolution of behaviour, and sexual isolation, concluding that "natural selection is opportunistic and operates upon whole animals not upon parts." The last two chapters on *Learning* deal with the classification of different types of learning, the comparative study of learning, and the physical basis of learning. While the classification of learning is difficult to relate to experimental operations and rather outdated, one would like to see sections particularly relevant to the previous chapters expanded (e.g. comparative learning expanded to give a further idea of just what sorts of comparisons are meaningful).

Although presenting such a variety of material on such a variety of animals, the author offers no superficial solutions. He provides clear descriptions of experimental procedures and points out complications in interpretation. Animal behaviour is presented as an intriguing subject requiring further investigation.

JOAN G. STEVENSON

Thinking and Reasoning. Edited by P. C. Wason and P. N. Johnson-Laird. Harmondsworth: Penguin Books Ltd. 1968. Pp. 431. 8s. 6d.

If Dr. Wason and Dr. Johnson-Laird are looking for more material for tests of logical abilities, they ought to try this statement: "All the experiments in our book of selected readings are about problem solving by human subjects. Our book is called 'Thinking and Reasoning'. Therefore human subjects only think and reason in problem solving experiments: they do not reason or even think in experiments on verbal learning, verbal recall, probability learning, perceptual constancy or vigilance." In fact it seems that, in a manner peculiarly appropriate to their own research interests, these editors have defined the subject matter of their book entirely by the exclusion of negative instances. The excluded negative instances are all those experiments in which some form of hypothesis formation is studied, but which nevertheless may be included in a more established category, such as human learning or perception. What is left is human problem solving, and since many students might rely on such a collection as a definition of the content of an area of research, "Human Problem Solving" is what this book should have been called.

If this collection is to be judged on how it manages to present an accurate and coherent picture of research on problem solving, then it must be judged, with one notable exception, to be a success. It has six sections, and three of these are excellent. These three concern Deductive and Inductive Reasoning and Matching Problems. Each section contains an account of a specific area of research in which reports of experimental progress are nicely balanced with discussion of specific problems of experimental design. Of the remaining three sections, two raise certain problems, although they both contain much that is interesting. One can question, for example, in the section on computer simulation why the ARGUS program was included and the EPAM model left out. In another section of which the title is actually "Problem Solving," one can detect signs that the editors are functionally fixed on the topic of functional fixedness. Four out of six papers make the well-worn point that experience can hinder original solutions to problems. One of these might have been abandoned in favour of Yarbus' account of how eye movement studies can be used to detect the effects of previous experience on the way information is assessed in a new situation. It is in the sixth section on development that the most striking imbalance is to be found. This section is devoted entirely to Piaget and his colleagues and to Bruner. Piaget's position cannot be questioned, although the selection of his work here is very odd. Only fleeting references are allowed to the classic conservation and seriation experiments, and a great deal of space is devoted to his more dubious and less replicable work on class inclusion. Bruner's place is more questionable, since most of his work in this field is derivative and he consistently fails to consider, let alone to eliminate, alternative hypotheses (suitable case, one would have thought, for Dr. Wason). What really is surprising here is the omission in this section and even in the suggestions for further reading of any mention of the major Russian work in this area, the work of Luria, Zaporhzets, Liublinskaya, Zinchenko and Lomov. Many accounts of these exist in English. Anyone interested in this section would do well also to look further afield.

Each section is preceded by a brief introduction, but as in most books of readings the editors give no account of what they actually think about the articles which they have chosen. It is my opinion that there should be a rule that editors should at least say whether they choose papers because they are good, or because they are representative, or both. This might help the reader. It would certainly make reviewers' lives a great deal easier.

P. E. BRYANT

Les Troubles de la Résolution de Problémes. By A. R. Luria and L. S. Tsvetkova. Paris: Gauthier-Villars. 1967. Pp. 271.

Neuropsychological Symptoms and Syndromes in Lesions of the Occipital Lobe and the Adjacent Areas. By I. Gloning, K. Gloning and H. Hoff. Paris: Gauthier-Villars. 1968. Pp. 98.

Performances in Aphasia. A Neurodynamical, Diagnostic and Psychological Study. By A. Kreindler and A. Fradis. Paris: Gauthier-Villars. 1968. Pp. 261.

These three volumes, together with a French translation of Hermelin and O'Connor's *Speech and Thought in Severe Subnormality*, are the first to appear in a new series *Collection Neuropsychologia: Monographies de Neuropsychologie* under the Editorship of Professor Henry Hécaen. The Luria-Tsvetkova book is concerned with the analysis of problem-solving defects in cases of frontal lesion and has been excellently translated by Dr. R. Lhermitte. The Glonings and Hoff provide a clinical and statistical analysis of 241 cases of posterior hemisphere lesions, verified anatomically either at operation or autopsy. Kreindeer and Fradis report new chemical and experimental studies in aphasia. This series should be of particular interest to experimental psychologists and others who have occasion to work in the neurological field.

O. L. ZANGWILL

Hearing Mechanisms in Vertebrates: A CIBA Foundation Symposium. Edited by A. V. S. de Rueck and Julie Knight. London: Churchill. 1968. Pp. xi + 320. 65s.

The publicity for this volume describes it as of interest to neurophysiologists, anatomists, neurologists and otologists. The poor psychologist is not invited. Nor was he, indeed, to the symposium that the book reports. This in itself is an interesting fact. Another recently published book on audition, Whitfield's *The Auditory Pathway*, conspicuously omits reference to psychological evidence. Do the physiologists no longer need psychological observations now they can get inside the system with micro-electrodes and are beginning to sort it out? Scientifically oriented psychologists have long favoured working in an area where the relation between results of behavioural experiments and plausible physiological mechanisms was reasonably close. Have they died out or been given the sack?

In fact they are alive and producing excellent material at a great rate. In a decade the Journal of the Acoustical Society of America has expanded threefold, and the psychoacousticians have more than held on to their share of the total. The understanding of the mechanisms of masking, frequency discrimination, localization, fatigue and loudness perception has increased considerably. The relationship between the psychological results and known or verifiable physiological mechanisms continues to be close. One is tempted to conclude that progress has been too fast for the physiologists to keep up. It might be safer to conclude that this is a predominantly European symposium and that psychoacoustics has not had such a big impact in Europe as in America. These are fairly uncontroversial facts.

But to the book. It contains fourteen presentations in five groups: Structure and Function of Hearing Organs in Non-mammalian Vertebrates; Structure and Function in Mammals; Afferents and Efferents in the Auditory Nerve; Spatial Localization; Central Mechanisms. The two most interesting papers are a neuro-anatomical one by Spoendlin and a neurophysiological one by Fex, on the microanatomy of the organ of corti, and on the role of the efferents to the cochlea, respectively. It appears we are now half way to knowing how the nerve cells are triggered by the hair cells, and what the bundles of efferent fibres do.

One of the most valuable parts of the book is the report of discussions. Typically with such small high-powered groups, these discussions contain a wealth of offbeat information that one vaguely knows about but cannot lay ones hands on—information that is problem-oriented rather than presentation-oriented.

The volume has an appealing aroma of high-quality paper and is full of photographs and diagrams. Despite its lack of psychoacoustics it will be a standard reference in hearing for many years to come. And rightly so.

MARK P. HAGGARD

Language: Selected Readings. Edited by R. C. Oldfield and J. C. Marshall. Harmondsworth: Penguin Books. 1968. Pp. 392. 8s. 6d.

As books of readings go, this is pretty good, certainly streets ahead of a similar selection previously evaluated by this reviewer. Each paper has a claim to fame and inclusion, although a different pair of editors would probably have omitted some and included others. The introductions are short and to the point. One would have preferred a text, but this book is definitely of use in advanced teaching.

MARK P. HAGGARD

Studies in Psychiatric Art: Its Psycho-Dynamics, Therapeutic Value and Relationship to Modern Art. By R. W. Pickford. Springfield, Illinois: Thomas. 1967. Pp. xix + 348. $14.00.

Professor Pickford has collected together a number of his articles and papers in the general field of art and psychopathology and added a certain amount of unpublished material. As might be expected from this author, the point of view is Freudian, though not narrowly so—indeed one chapter betrays considerable sympathy with Jung. The result, though somewhat lacking in unity and theme, is an interesting and in patches extremely perceptive study.

O. L. ZANGWILL

Aggression and Defense. Edited by C. D. Clemente and D. B. Lindsley. Los Angeles: University of California Press. London: Cambridge University Press. 1968. Pp. xv + 361. £7 2s. 6d.

This book is the proceedings of a conference held in California in November, 1965, with participants from a wide range of disciplines. The initial similarity between this and an earlier book, Carthy and Ebling (1964), proves deceptive, and if the reader is not deterred by the frontispiece he may find many of the chapters useful for the data they contain or as reviews of areas of research. Other chapters will be much less useful. The increasing frequency with which books such as this are produced raises the question of whether they have a merit other than that of their constituent papers. I think that they may have where the discussions show which are superficial and which are basic points of disagreement between authors with apparently opposing views, e.g. I. Eibl-Eibesfeldt and J. P. Scott in this volume.

In the opening paper, L. S. B. Leakey suggests that weapons were first used by pre-man when he turned to a scavenging existence, to drive other scavengers from the kill: later they were used to kill prey, and by a natural progression, to kill his fellows. This is followed by papers on animal behaviour and the physiology of agressive behaviour. The Cetacea are discussed by K. S. Norris. The widespread occurrence of ritualized fighting in animals which possess dangerous weapons suggests that strong selection pressures operate to retain the use of aggressive behaviour as a spacing out mechanism, and I. Eibl-Eibesfeldt and S. A. Barnett discuss the way in which species resolve conflicts without the use of harmful weapons. Neither author attempts a comprehensive review of the literature on animal fighting, and there is little mention of fish behaviour, where much of the quantitative work has been performed, or of the relationship between courtship and fighting displays. Eibl-Eibesfeldt also deals with the effect of isolation on aggressive behaviour, and this section shows clearly how his (Lorenzian) view of aggression differs from that of e.g. J. P. Scott. Reaction to a strange animal with the species-specific motor patterns of attack does not require the previous experience of conspecifics: hence fighting is not solely learned behaviour. Studies showing that the opportunity to attack is reinforcing are interpreted as evidence of an "urge to fight." Finally, examples of fluctuations in the threshold for attack are discussed. If a male Cichlid *Etroplus maculatus* could attack other males through a glass plate, few attacks were made on the mate, while if the pair were visually isolated, the mate was attacked and killed. Eibl-Eibesfeldt suggests that this is the result of damming up the male's aggression. This increase in responsiveness is the vital point in his argument for spontaneity, and it is not clear whether it is due to the removal of a short term inhibition, or whether it is a long term, continuous threshold lowering. Heiligenberg's work on *Pelmatochromis*, also quoted, shows that the lowering of the threshold for aggressive behaviour is not an invariable consequence of isolation in Cichlids, and it seems important that these increases in responsiveness with time should be investigated by those who hold that aggression is not spontaneous, as well as by those who hold that it is.

S. Kawamura shows how complex aggressive behaviour may be in troops of Japanese macaques. Thus, a monkey may attack an inoffensive bystander to divert an attack from itself (by inducing its own attacker to join in this attack), or to apply psychological pressure to a leader (by demonstrating that it lacks the superiority to punish this attack as its status demands it should). By joining in and leading the attacks of a dominant animal, a monkey may obtain the other's friendship, and hence raise its own status. Moreover, there are differences between troops not only in tolerance of proximity, but in the possession of particular ritualized behaviour patterns.

This chapter shows clearly how inadequate an analysis of attacks in simple stimulus-response terms may be. A chapter by J. M. R. Delgado shows how by radiostimulating Rhesus monkeys with chronically implanted electrodes in a social group, it may be possible to study such complex behaviour physiologically. Brain loci at which stimuli evoked stereotyped gestures and grimaces which might appear aggressive, but which were not socially directed, could be separated from those producing, as an after-effect of stimulation, well co-ordinated attacks of variable form directed animals with which relations were normally unfriendly. The contribution of more conventional ablation and stimulation studies to the problem of the neural organization of aggressive behaviour is reviewed at length by B. Kaada. A. B. Rothballer eschews psychopharmacology and adopts an ecological approach when discussing the relations between neurohumours (which he defines broadly) and aggressive behaviour. Interspecific differences must be borne in mind in a search for the physiological correlates of aggressive behaviour. Rothballer cites evidence to show that while in species in which territory formation and breeding occur together, testosterone controls the aggressive behaviour, LH may be important in species in which territories are claimed before much testis growth has occurred: he does not point out that the experiments in this area urgently need repeating and extending. Changes in hormone levels consequent on aggressive behaviour are then discussed, and B. L. Welch adds a comment on the relations between catecholamines and fighting.

The four final papers, on human behaviour, differ widely. L. Berkowitz reviews a small body of data cues provided by a target associated with aggression in the individual's past which increases the aggression elicited. Cognitive processes may be modified in an attempt to provide a justification for these responses. H. W. Brosin provides a historical review of ways in which aggression is controlled, from a psychiatrist's point of view, and the book ends with chapters by S. C. Plog on the methods used to study the Watts riots of 1965, and by H. D. Lasswell on the social and political framework of war and peace. There are subject and name indices.

For the reader who wants a handy source of references to a restricted field, the wide scope of this book will be a disadvantage, but the wide range of views expressed is welcome in the light of the one-sided theses propounded in some recent, and very popular, books on aggression.

P. G. CARYL

Human Aggression. By Anthony Storr. London: Allen Lane The Penguin Press. 1968. Pp. vi + 127. 25s.

This book is evidently intended for the naïve reader—in the sense in which experimentalists refer to "naïve rats"—that is, it is essentially for those who have not already been exposed to theories of aggression, for it does not have a great deal to say that is new. It is an account of human aggression, as understood by the psycho-analysts. First it is compared with aggressive behaviour in the lower animals; it is then discussed with respect to infants, children and adults, and then with respect to sexual relationships. In the later chapters, aggression is discussed in the depressive, the schizoid, the paranoid and the psychopath. The last chapter is on "Ways of reducing hostility."

The main features are the repeated stress (i) on the long period of development required by the human being before he attains physical and emotional maturity: the dependence and vulnerability engendered by this is thought to be one of the reasons underlying the wanton aggression and cruelty which is found in *homo sapiens* alone. (ii) Dr. Storr is at pains to point out that aggression has good features as well as deplorable ones: he evidently considers it as potent a force as the sexual impulse in artistic and intellectual creativity. This leads to (iii) the impression given by Dr. Storr that women constitute an inferior minority of the population whose only claim to creativity is in the strictly biological field. That this is no overstatement may be seen in his consolatory message to womankind: "But it is equally vital that we should be cherished and fed, and that we should reproduce ourselves. Women have no need to compete with men; for what they alone can do is the more essential."

This does not seem wholly compatible with the (readily acceptable) statement in the final chapter that the reduction of world population "is the most important single step which can be taken by mankind to reduce hostile tension." This chapter contains some other practical suggestions for decreasing international tension, but it neglects to mention the substantial financial gain made by a few as a result of war.

A. W. HEIM

A Textbook of Physiological Psychology. By Sebastian P. Grossman. London and New York: Wiley. 1967. Pp. xix + 932. 113s.

Dr. Grossman has written an excellent comprehensive text for physiological psychology which deserves a place on the reference shelves of all psychology libraries. It provides a wide and accurate introduction and review of almost all the major aspects of physiological psychology. It should be useful to lecturers and research workers and should certainly be strongly recommended to students with physiological interests.

The text is divided into four parts, Part I (174 pages) consists of the biophysics, basic electrophysiology and anatomy of the nervous system. Part II (132 pages) is concerned with sensory processing, sensory-motor integration and the reticular activating system. Part III (285 pages) describes a wide range of experimental work on motivational mechanisms. Part IV (258 pages) is concerned with the plasticity of the organism's response repertoire and mechanisms of learning.

The depth of coverage of these subjects varies greatly. For example, Part III on motivational mechanisms includes extremely comprehensive chapters on Hunger, Thirst, Sexual Behaviour, Emotional Behaviour and Self-stimulation. This after all is Dr.

Grossman's own field of research and these chapters will undoubtedly be of great value to students and research workers alike. It seems a pity that the author's personal views of motivational mechanisms were not more apparent as he has been in the forefront of this field for a number of years.

On the other hand, Chapter 5 on the reticular formation is very limited in its coverage. The mechanisms of sleep and paradoxical sleep are barely mentioned, which is a sad omission in view of the enormous current interest in these subjects.

It would obviously be impossible in one text to give detailed coverage to all the subjects listed in the index and possibly it would have been wise to use greater selectivity. For example the first 177 pages describe the basic biophysics, electrophysiology and anatomy of the nervous system in more detail than is perhaps necessary, especially in view of the fact that excellent physiology and anatomy texts exist to provide the student with this information. Also, it seems difficult to justify the inclusion of part of Chapter 3 on vision, which takes the reader through the most elementary psychophysics and description of visual illusions before discussing the electrophysiology of the visual system. On the other hand, the auditory and somaesthetic systems are not described in any detail and perhaps a more general appraisal of the sensory processing mechanisms would have produced a more global and relevant approach to the senses for the physiological psychologist.

From a different point of view, it is unfortunate that the final chapter on biochemical methods and learning is not more critical. It is admirable that this subject is introduced with some basic biochemical facts about nucleic acids and their synthesis but unfortunate that the controversial experiments themselves, which attempt to relate changes in nucleic acid synthesis with learning and memory were presented without, what many consider, the very necessary caution.

It is debatable what function a textbook fulfils, but in a subject like physiological psychology it should be more than a statement of the facts from the contributory disciplines. This textbook is presumably aimed at final year undergraduates and postgraduates and considering its audience a little more interpretation and synthesis by the author would have made the enormous amount of experimental data presented both more palatable and exciting.

Nevertheless, Dr. Grossman has presented a massive amount of material accurately, the bibliography is extensive and will be very useful to student, lecturer and researcher. The text is well produced and the information is arranged clearly under numerous headings and sub-headings. The photographic work is of a high standard with informative captions. The usefulness of the illustrations is marred only by the failure in the written text to refer by number or letter to the diagrams. This is particularly noticeable in Part III of the book, which is in many other ways excellent.

S. D. IVERSEN

Q. Jl exp. Psychol. (1969) **21**, i–iii

PROCEEDINGS OF THE EXPERIMENTAL PSYCHOLOGY SOCIETY, 1968

28th–29th March 1968. Meeting at Cambridge.

1st Session: "Some experiments on simultaneous translation," by Mr. D. Gerver.† "Delayed auditory feedback: a limited capacity model," by Mr. D. Salter.† "Temporal uncertainty in choice-reaction experiments," by Dr. D. R. J. Laming.

2nd Session: "Disorders of visual perception in patients with localized cerebral lesions," by Dr. Elizabeth Warrington. "An automated technique for the study of learning ability in psychogeriatric patients," by Dr. A. W. H. Buffery, Mr. C. J. Cheney† and Dr. J. L. Gedye. "Speech and memory during anaesthetization of one or other cerebral hemisphere," by Dr. Brenda Milner.

3rd Session: "Maintenance of responding by a fixed-interval schedule of electric-shock presentation in squirrel monkeys (*Saimiri sciurens*)," by Dr. R. G. A. Stretch. "A new theory of visual spatial resolution based on neuro-physiological findings," by Dr. F. W. Campbell, Dr. G. F. Cooper† and Dr. J. G. Robson.†

4th Session: "On the measurement of loudness," by Professor M. Treisman. "The invariance of critical bandwidth," by Dr. M. P. Haggard. "Massed and distributed practice—is it a dead issue?" by Dr. A. D. Baddeley and Miss D. J. A. Longman.† "Anxiety stress and decrements in skill," by Dr. M. Hammerton. "Interaction of verbal instructions and task difficulty in the learning of perceptual-motor skill," by Mr. B. R. Gaines.†

16th–17th July 1968. Meeting at Nottingham.

1st Session: "Problems of measuring performance in blind mobility," by Dr. J. A. Leonard. "The heart-rate of blind and sighted pedestrians on an unfamiliar route," by Mr. R. Wycherley† and Miss B. Nicklin.† "Physiological measures of stress," by Dr. M. Hammerton. "Baseline determinants of conditioned suppression, comparing two research strategies," by Dr. D. E. Blackman.

2nd Session: "Topological representations and the structure of verbal problems," by Mr. D. Woods.† "Dissociated deficits in right hemisphere lesions," by Dr. Freda Newcombe.† "The synaptic basis of memory," by Dr. J. A. Deutsch.

3rd Session: "Pre-perceptual acoustic storage," by Dr. J. Morton and Dr. R. G. Crowder.† "Recognition and number of choices: A-index and d′ comparisons," by Professor J. Brown. "A decision theory approach to signal detection and its application to detection latencies: some further developments," by Dr. G. Wolfendale.†

4th Session: "Interaction of visual stimuli: Benham's top and Bidwell's disc," by Dr. A. H. Gregory. "Experiments on timing in natural speech," by Dr. A. W. F. Huggins.† "Speed-error trade-off in serial choice reaction tasks," by Mr. D. Hale.† "Temporal uncertainty in choice-reaction experiments—2", by Dr. D. R. J. Laming. "Camouflage and visual search," by Professor C. I. Howarth and Mr. J. R. Bloomfield.† "Stimulus uncertainty in visual detection and recognition," by Dr. M. Matthews† and Mr. J. Kemp.†

† By invitation

6th–7th January 1969. Annual General Meeting at Bedford College, London.

1st Session: "Serial and parallel processing in pattern recognition," by Mr. A. J. Marcel.† "Some problems in the hemispheric control of reaction time," by Dr. S. J. Dimond.† "Attentional narrowing in loud noise," by Mr. G. R. J. Hockey.† "A neurophysiological model of beta apparent movement and some relevant psychological data," by Mr. J. P. Frisby.†

2nd Session: "Intra-trial interference in short-term memory," by Dr. H. Loess. "Specific impairment of auditory verbal short-term memory," by Dr. Elizabeth Warrington and Dr. T. Shallice. "The assessment and analysis of handedness," by Professor R. C. Oldfield.

3rd Session: "Adaptation of divers to size-distance distortion under water," by Dr. Helen E. Ross, Mr. S. S. Fanklin† and Mr. G. Weltman.† "Some experiments on pitch perception and their relation to frequency control in the human voice," by Mr. A. Rostron.† "Effects of frontal cortex lesions on the development of visual discrimination learning set in the monkey," by Dr. J. M. Oxbury. "Inter-ocular transfer in the hooded rat," by Dr. L. Nadel.†

4th Session: "The role of the cerebral cortex in passive avoidance learning," by Mr. D. Kleinman.† "The hemidecorticate learning deficit," by Dr. H. C. Plotkin.† "A parametric study of lateralization in the functional split-brain," by Dr. I. S. Russell. "On the relationship between letter names and superordinate categories," by Dr. M. I. Posner. "The effect of task difficulty on the relative speeds of correct and incorrect responses in an identification task," by Mr. J. Wilding.†

† By invitation

Committee, 1969: *President:* Professor P. H. Venables.

Editor: Mr. A. J. Watson.

Ordinary Members of the Committee: Dr. J. Annett, Dr. S. Anstis, Dr. A. D. Baddeley (Hon. Treasurer), Dr. D. Legge (Hon. Secretary), Dr. N. Moray, Dr. P. M. A. Rabbitt (Hon. Ass. Secretary), Dr. I. S. Russell, Dr. A. W. Still, Dr. Elizabeth Warrington.

The following have been elected into membership of the Society:

D. A. Allport, Ph.D., Department of Psychology, Kings College, Old Aberdeen.
C. B. Blakemore, Ph.D., Physiological Laboratory, Downing Street, Cambridge.
I. D. Brown, Ph.D., M.R.C., A.P.R.U., 15 Chaucer Road, Cambridge.
B. Craske, Ph.D., Department of Psychology, The University, Southampton.
J. M. Cullen, Ph.D., Institute of Experimental Psychology, 1 South Parks Road, Oxford.
D. R. Davies, Ph.D., Department of Psychology, The University, Leicester.
I. E. Gordon, Ph.D., Department of Psychology, University of Exeter.
N. K. Humphrey, Ph.D., Institute of Experimental Psychology, 1 South Parks Road, Oxford.
D. Jones, Ph.D., Department of Psychology, Birkbeck College, Malet Street, London, W.C.1.
G. R. Kiss, B.Sc., M.R.C. Speech and Communication Research Unit, University of Edinburgh, 31 Buccleuch Place, Edinburgh 8.
Mrs. Freda Newcombe, D.Phil., Neuropsychology Department, Churchill Hospital, Headington, Oxford.
K. G. Oatley, Ph.D., Laboratory of Experimental Psychology, University of Sussex, Falmer, Brighton, BN1 9QY.
E. A. Salzen, Ph.D., Department of Psychology, Kings College, Old Aberdeen.

M. H. Sheldon, Ph.D., Department of Psychology, The University, Whiteknights, Reading.
Miss J. G. Stevenson, Ph.D., Psychological Laboratory, Downing Street, Cambridge.
E. A. C. Thomas, Ph.D., Department of Psychology, University College, Gower Street, London, W.C.1.

Visiting Foreign Members

I. D. John, Ph.D., Department of Psychology, The University, Reading.
H. Loess, Ph.D., M.R.C., A.P.R.U., 15 Chaucer Road, Cambridge.
D. Lykken, Ph.D., Department of Psychology, Birkbeck College, Malet Street, London, W.C.1.
A. Mechanic, Ph.D., Department of Psychology, University College, Gower Street, London, W.C.1.
J. M. Millensen, Ph.D., Department of Psychology, The University, Whiteknights, Reading.
J. Nachmias, Ph.D., Psychological Laboratory, Downing Street, Cambridge.
M. I. Posner, Ph.D., M.R.C., A.P.R.U., 15 Chaucer Road, Cambridge.
A. H. Winefield, Ph.D., Visiting Great Britain from Department of Psychology, University of Adelaide, South Australia.

Q. Jl. exp. Psychol. (1969) **21,** iv–ix.

MEMBERS OF THE EXPERIMENTAL PSYCHOLOGY SOCIETY

Honorary Member

Sir Frederic Bartlett, C.B.E., F.R.S., St. John's College, Cambridge.

Ordinary Members

D. A. Allport, Ph.D., Department of Psychology, Kings College, Old Aberdeen.
J. Annett, D.Phil., Department of Psychology, The University, Hull.
S. M. Anstis, Ph.D., Department of Psychology, The University, Bristol 8.
H. Asher, B.A., Department of Physiology, University of Birmingham, Birmingham 15.
Professor R. J. Audley, Ph.D., Department of Psychology, University College, Gower Street, London, W.C.1.
B. Babington Smith, O.B.E., M.A., Institute of Experimental Psychology, 1 South Parks Road, Oxford.
A. D. Baddeley, Ph.D., Laboratory of Experimental Psychology, University of Sussex, Falmer, Brighton, Sussex.
Professor P. Bakan, Ph.D., Department of Psychology, Michigan State University, East Lansing, Michigan, U.S.A.
Professor H. B. Barlow, M.D., M.B., B.Chir., University of California, School of Optometry, Berkeley 4, California, U.S.A.
Mrs. E. Belbin, Ph.D., 8 Lansdowne Road, Cambridge.
A. J. Benson, M.B., Ch.B., R.A.F. Institute of Aviation Medicine, Farnborough, Hants.
Professor D. E. Berlyne, Ph.D., Department of Psychology, University of Toronto, Toronto 5, Ontario, Canada.
Professor P. Bertelson, Ph.D., Laboratoire de Psychologie, Universite Libre de Bruxelles, 115 Avenue Adolphe Buyl, Bruxelles, Belgium.
D. E. Blackman, Ph.D., Department of Psychology, The University, Nottingham.
C. B. Blakemore, Ph.D., Physiological Laboratory, Downing Street, Cambridge.
T. M. Bloomfield, Ph.D., Laboratory of Experimental Psychology, University of Sussex, Falmer, Brighton.
D. A. Booth, Ph.D., Laboratory of Experimental Psychology, University of Sussex, Falmer, Brighton.
G. S. Brindley, M.D., F.R.S., Trinity College, Cambridge.
D. E. Broadbent, Sc.D., Applied Psychology Research Unit, 15 Chaucer Road, Cambridge.
I. D. Brown, Ph.D., Applied Psychology Research Unit, 15 Chaucer Road, Cambridge.
Professor J. Brown, Ph.D., Department of Psychology, The University, Bristol 8.
D. J. Bruce, Ph.D., Department of Education, 5 Salisbury Villas, Station Road, Cambridge.
P. E. Bryant, Ph.D., Institute of Experimental Psychology, 1 South Parks Road, Oxford.
A. W. H. Buffery, Ph.D., Human Development Research Unit, Park Hospital for Children, Old Road, Headington, Oxford.
F. W. Campbell, M.B., Ch.B., Ph.D., D.O.M.S., Physiological Laboratory, Downing Street, Cambridge.
Miss V. R. Cane, M.A., 14 Little St. Mary's Lane, Cambridge.
A. Carpenter, M.B., B.Chir., Applied Psychology Research Unit, 15 Chaucer Road, Cambridge.
Professor E. C. Cherry, D.Sc., Department of Electrical Engineering, Imperial College of Science and Technology, Exhibition Road, London, S.W.7.

M. B. Clowes, Ph.D., C.S.I.R.O., Computing Research Section, P.O. Box 109, Canberra City, A.C.T., Australia.
W. P. Colquhoun, Ph.D., Applied Psychology Research Unit, 15 Chaucer Road, Cambridge.
K. J. Connolly, B.Sc., Department of Psychology, The University, Sheffield 10.
R. Conrad, Ph.D., Applied Psychology Research Unit, 15 Chaucer Road, Cambridge.
D. W. J. Corcoran, Ph.D., Applied Psychology Research Unit, 15 Chaucer Road, Cambridge.
A. Cowey, Ph.D., Institute of Experimental Psychology, 1 South Parks Road, Oxford.
J. J. Cowley, Ph.D., Department of Psychology, Queen's University of Belfast, 1 Lennoxvale, Belfast 7.
F. I. M. Craik, Ph.D., Department of Psychology, Birkbeck College, Malet Street, London, W.C.1.
B. Craske, Ph.D., Department of Psychology, The University, Southampton.
Miss A. Crawford, Ph.D., Department of Psychology, The University, 7 Abercrombie Square, Liverpool 7.
Professor E. R. F. W. Crossman, Ph.D., Department of Industrial Engineering, University of California, Berkeley 4, California, U.S.A.
S. Crown, Ph.D., 5 Northwick Avenue, Kenton, Middlesex.
J. M. Cullen, Ph.D., Institute of Experimental Psychology, 1 South Parks Road, Oxford.
H. C. A. Dale, Ph.D., Department of Psychology, The University, Hull.
Professor K. Danziger, D.Phil., Department of Psychology, York University, Keele Street, Toronto 12, Ontario, Canada.
D. R. Davies, Ph.D., Department of Psychology, The University, Leicester.
R. Davis, D.Phil., Department of Psychology, The University, Reading, Berks.
B. Delisle Burns, M.R.C.S., L.R.C.P., Division of Physiology and Pharmacology, National Institute for Medical Research, Mill Hill, London, N.W.7.
Professor J. A. Deutsch, D.Phil., Department of Psychology, University of California, San Diego, La Jolla, California, U.S.A.
N. F. Dixon, Ph.D., Department of Psychology, University College, Gower Street, London, W.C.1.
P. C. Dodwell, Ph.D., Department of Psychology, Queen's University, Kingston, Ontario, Canada.
J. Drever, M.A., Principal, University of Dundee, College Gate, Dundee, Scotland.
Professor G. C. Drew, M.A., Department of Psychology, University College, Gower Street, London, W.C.1.
Professor J. T. Eayrs, D.Sc., Department of Anatomy, The Medical School, Edgbaston, Birmingham 15.
A. Elithorn, M.D., M.B., B.Chir., M.R.C.P., D.P.M., Royal Free Hospital, Lawn Road, Hampstead, London, N.W.3.
G. Ettlinger, Ph.D., Institute of Psychiatry, De Crespigny Park, Denmark Hill, London S.E.5.
Professor B. M. Foss, M.A., Department of Psychology, Bedford College, Regent's Park, N.W.1.
Professor E. Fraser, Ph.D., Department of Psychology, The University, Aberdeen.
J. L. Gedye, M.B., B.Chir., Department of Electrical Engineering Science, University of Essex, Colchester, Essex.
F. H. George, Ph.D., Department of Psychology, The University, Bristol 8.
K. Gibbins, Ph.D., Department of Psychology, The University, 7 Sydenham Terrace, Newcastle upon Tyne 2.
Mrs. F. Goldman-Eisler, D.Phil., Psycholinguistics Research Unit, Department of Phonetics, University College, Gower Street, London, W.C.1.
I. E. Gordon, Ph.D., Department of Psychology, University of Exeter, Exeter.
G. W. Granger, Ph.D., Department of Psychology, 7 Kepier Terrace, Gilesgate, Durham.
J. A. Gray, Ph.D., Institute of Experimental Psychology, 1 South Parks Road, Oxford.

A. H. Gregory, Ph.D., Department of Psychology, The University, Manchester 13.
Professor R. L. Gregory, M.A., Department of Machine Intelligence and Perception, Bionics Research Laboratory, The University, Forrest Road, Edinburgh 1.
Professor R. A. M. Gregson, Ph.D., Department of Psychology and Sociology, University of Canterbury, P.O. Box 1471, Christchurch, C.I., New Zealand.
Professor S. Griew, Ph.D., Department of Psychology, The University, Dundee.
G. C. Grindley, M.A., Psychological Laboratory, Downing Street, Cambridge.
M. P. Haggard, Ph.D., Psychological Laboratory, Downing Street, Cambridge.
A. M. Halliday, M.B., Ch.B., M.R.C., Neurological Unit, National Hospital, Queen Square, London, W.C.1.
M. S. Halliday, Ph.D., Laboratory of Experimental Psychology, University of Sussex, Falmer, Brighton, Sussex.
M. Hammerton, Ph.D., Applied Psychology Research Unit, 15 Chaucer Road, Cambridge.
R. Harper, Ph.D., 18 Leng Crescent, Norwich.
A. D. Harris, M.R.C.S., L.R.C.P., Littlemore Hospital, Littlemore, nr. Oxford.
G. J. Harrison, Ph.D., Department of Psychology, The University, Sheffield, S10 2TN.
Miss A. W. Heim, Ph.D., Psychological Laboratory, Downing Street, Cambridge.
L. J. Herberg, M.B., B.Ch., D.Phil., Institute of Neurology, National Hospital, Queen Square, London, W.C.1.
Mrs. B. M. Hermelin, Ph.D., Developmental Psychology Research Unit, Drayton House, Gordon Street, W.C.1.
W. E. Hick, M.D., M.B., 66 Bishop's Road, Trumpington, Cambridge.
R. Hinchcliffe, M.D., Institute of Laryngology and Otology, Royal National Hostpital for Throat, Nose and Ear, Grays Inn Road, London, W.C.1.
Professor R. A. Hinde, D.Phil., Sub-Department of Animal Behaviour, High Street, Madingley, Cambridge.
D. H. Holding, Ph.D., Department of Psychology, The University, Gandy Street, Exeter.
G. Horn, M.B., Ch.B., School of Anatomy, The University, Cambridge.
N. Hotoph, M.A., London School of Economics, Houghton Street, Aldwych, London, W.C.2.
I. P. Howard, Ph.D., Department of Psychology, York University, Keele Street, Toronto 12, Ontario, Canada.
Professor C. I. Howarth, D.Phil., Department of Psychology, The University, Nottingham.
J. A. M. Howe, Ph.D., Department of Machine Intelligence and Perception, Bionics Research Laboratory, The University, Forrest Road, Edinburgh 1.
D. H. Howes, Ph.D., Room E6–121 (Neurology), Boston Veterans Administrative Hospital, 150 South Huntington Avenue, Boston, Massachusetts 02130, U.S.A.
I. M. Hughes, Ph.D., Department of Psychology, Taylor Building, The University, Aberdeen.
N. K. Humphrey, Ph.D., Institute of Experimental Psychology, 1 South Parks Road, Oxford.
Professor I. M. L. Hunter, D.Phil., Department of Psychology, The University, Keele, Staffs.
Professor H. M. B. Hurwitz, Ph.D., Department of Psychology, University of Tennessee, Knoxville, Tennessee 37916, U.S.A.
Miss M. Ikeda, Ph.D., Research Department of Ophthalmology, Royal College of Surgeons of England, Lincoln's Inn Fields, London, W.C.2.
Mrs. S. D. Iversen, Ph.D., Psychological Laboratory, Downing Street, Cambridge.
Professor P. H. R. James, Ph.D., Department of Psychology, Dalhousie University, Halifax, Nova Scotia, Canada.
Professor M. A. Jeeves, Ph.D., Department of Psychology, University of Adelaide, G.P.O. Box 198 D., Adelaide, S. Australia.
P. N. Johnson-Laird, Ph.D., Department of Psychology, University College, Gower Street, London, W.C.1.
A. R. Jonckheere, Ph.D., Department of Psychology, University College, Gower Street, London, W.C.1.

D. Jones, Ph.D., Department of Psychology, Birkbeck College, Malet Street, London, W.C.1.
Mrs. D. Joyce, Ph.D., Department of Psychology, Birkbeck College, Malet Street, London, W.C.1.
Professor H. Kay, Ph.D., Department of Psychology, The University, Sheffield 10.
R. P. Kelvin, Ph.D., Department of Psychology, University College, Gower Street, London, W.C.1.
Mrs. M. O. Kerr, B.A., Royal Free Hospital, North Western Branch, Lawn Road, London, N.W.3.
C. C. Kiernan, B.A., Department of Psychology, Birkbeck College, Malet Street, London, W.C.1.
M. Kinsbourne, B.M., B.Ch., M.R.C.P., 44 Fellows Road, London, N.W.3.
G. R. Kiss, B.Sc., M.R.C., Speech and Communication Research Unit, University of Edinburgh, 31 Buccleuch Place, Edinburgh 8.
D. R. J. Laming, Ph.D., Psychological Laboratory, Downing Street, Cambridge.
D. Legge, Ph.D., Department of Psychology, University College, Gower Street, London, W.C.1.
J. A. Leonard, Ph.D., Department of Psychology, The University, Nottingham.
B. N. Lewis, B.Sc., 20, Hill Rise, Richmond, Surrey.
N. E. Loveless, Ph.D., Department of Psychology, St. Andrew's University, Queen's College, Dundee.
D. J. McFarland, D.Phil., Institute of Experimental Psychology, 1 South Parks Road, Oxford.
Professor D. M. MacKay, Ph.D., Department of Communications, The University, Keele, Staffs.
N. J. S. M. Mackintosh, Ph.D., Department of Psychology, Dalhousie University, Halifax, Nova Scotia, Canada.
N. H. Mackworth, M.B., Ch.B., Ph.D., 429 Atlantic Street, Stamford, Conn., U.S.A.
Miss F. E. MacNeill, Ph.D., Department of Psychology, The University, Oxford Road, Manchester 13.
E. M. Macphail, D.Phil., Laboratory of Experimental Psychology, University of Sussex, Falmer, Brighton.
J. C. Marshall, Ph.D., M.R.C. Speech and Communication Research Unit, The University, 31 Buccleuch Place, Edinburgh 8.
Mrs. I. Martin, Ph.D., Institute of Psychiatry, De Crespigny Park, Denmark Hill, London, S.E.5.
Miss S. Mayne, Ph.D., 88, Bulmershe Road, Reading, Berks.
R. Melzack, Ph.D., Department of Psychology, McGill University, Montreal, Quebec, Canada.
Professor T. R. Miles, M.A., Department of Psychology, University College of North Wales, Bangor.
Mrs. B. Milner, Ph.D., Montreal Neurological Institute, 3801 University Street, Montreal 2, Quebec, Canada.
N. Moray, Ph.D., Department of Psychology, The University, Sheffield 10.
J. Morton, Ph.D., Applied Psychology Research Unit, 15 Chaucer Road, Cambridge.
G. H. Mowbray, Ph.D., Applied Physics Laboratory, Johns Hopkins University, Silver Spring, Maryland, U.S.A.
N. Mrosovsky, Ph.D., Department of Zoology, Toronto University, Toronto 5, Ontario, Canada.
W. R. A. Muntz, D.Phil., Laboratory of Experimental Psychology, University of Sussex, Falmer, Brighton, Sussex.
D. J. Murray, Ph.D., Department of Psychology, Queen's University, Kingston, Ontario, Canada.
Mrs. Freda Newcombe, D.Phil., Neuropsychology Department, Churchill Hospital, Headington, Oxford.

L. J. Newson, Ph.D., Department of Psychology, The University, Nottingham.

K. G. Oatley, Ph.D., Laboratory of Experimental Psychology, University of Sussex, Falmer, Brighton, BN1 9QY.

N. O'Connor, Ph.D., Developmental Psychology Research Unit, Drayton House, Gordon Street, London, W.C.1.

R. C. Oldfield, M.A., M.R.C. Speech and Communication Research Unit, The University, 31 Buccleuch Place, Edinburgh 8.

I. Oswald, M.D., D.Sc., D.P.M., University Department of Psychiatry, Royal Edinburgh Hospital, Morningside Park, Edinburgh 10.

J. M. Oxbury, Ph.D., Neurology Department, Churchill Hospital, Oxford.

S. Papert, Ph.D., Project M.A.C.–Room 816, Massachusetts Institute of Technology, Cambridge 39, Mass., U.S.A.

Miss M. Pickersgill, Ph.D., Department of Psychology, Bedford College, Regent's Park, London, N.W.1.

Professor R. W. Pickford, D.Litt., Ph.D., Department of Psychology, University of Glasgow, Glasgow, W.2.

M. F. Piercy, M.A., Maida Vale Hostpial for Nervous Diseases, Edgware Road, London, W.9.

I. Pleydell-Pearce, Ph.D., Department of Psychology, The University, Bristol 8.

E. C. Poulton, M.B., B.Chir., M.R.C.S., L.R.C.P., Applied Psychology Research Unit, 15 Chaucer Road, Cambridge.

Professor K. A. Provins, Ph.D., Department of Psychology, Australian National University, Box 4 P.O., Canberra, A.C.T. 2600, Australia.

P. M. A. Rabbitt, Ph.D., Institute of Experimental Psychology, 1 South Parks Road, Oxford.

Miss H. E. Ross, Ph.D., Department of Psychology, The University, Stirling.

W. A. H. Rushton, Ph.D., Sc.D., M.R.C.S., L.R.C.P., F.R.S., Trinity College, Cambridge.

I. S. Russell, Ph.D., Department of Psychology, University College, Gower Street, London, W.C.1.

Professor R. W. Russell, Ph.D., D.Sc., University of California, Irvine, Cal., U.S.A.

Professor D. Russell Davis, M.D., M.B., B.Chir., M.R.C.P., M.R.C.S., L.R.C.P., D.P.M., 39–41 St. Michael's Hill, Bristol, BS2 8DZ.

E. A. Salzen, Ph.D., Department of Psychology, King's College, Old Aberdeen.

B. Semeonoff, Ph.D., Department of Psychology, The University, 60 The Pleasance, Edinburgh 8.

L. M. Shaffer, Ph.D., Psychological Laboratory, The University, Gandy Street, Exeter.

T. Shallice, Ph.D., Department of Psychology, University College, Gower Street, London, W.C.1.

M. H. Sheldon, Ph.D., Department of Psychology, The University, Whiteknights, Reading.

M. E. Sime, B.A., Department of Psychology, The University, Sheffield 10.

Professor W. Sluckin, Ph.D., Department of Psychology, The University, Leicester.

J. R. Smythies, M.Sc., M.D., M.B., B.Chir., M.R.C.P., D.P.M., University Department of Psychiatry, Royal Edinburgh Hospital, Morningside Park, Edinburgh 10.

A. C. Staniland, Ph.D., Department of Psychology, The University, Gandy Street, Exeter.

Miss H. Steinberg, Ph.D., Department of Pharmacology, University College, Gower Street, London, W.C.1.

Miss J. G. Stevenson, Ph.D., Psychological Laboratory, Downing Street, Cambridge.

R. G. A. Stretch, Ph.D., Department of Psychology, Middlesex College, University of Western Ontario, London, Ontario, Canada.

A. W. Still, Ph.D., Department of Psychology, The University, 7 Kepier Terrace, Gilesgate, Durham.

Professor A. Summerfield, B.Sc., Department of Psychology, Birkbeck College, Malet Street, London, W.C.1.

Professor N. S. Sutherland, D.Phil., Laboratory of Experimental Psychology, University of Sussex, Falmer, Brighton, Sussex.

Professor J. R. SYMONDS, Ph.D., Psychology Department, King's College, Old Aberdeen.
Professor J. SZAFRAN, Ph.D., Department of Psychology, University of Southern California, University Park, Los Angeles, California 90007, U.S.A.
Professor H. TAJFEL, M.A., Department of Psychology, The University, Bristol 8.
D. H. TAYLOR, Ph.D., Psychology Department, The University, Southampton.
W. B. TEMPLETON, M.Sc, Department of Psychology, The University, 7 Kepier Terrace, Gilesgate, Durham.
E. A. C. THOMAS, Ph.D., Department of Psychology, University College, Gower Street, London, W.C.1.
Mrs. B. TIZARD, Ph.D., Department of Experimental Neurology, Institute of Psychiatry, De Crespigny Park, Denmark Hill. London, S.E.5.
Mrs. A. TREISMAN, D.Phil., Institute of Experimental Psychology, 1 South Parks Road, Oxford.
Professor M. TREISMAN, M.B., B.Ch., Department of Psychology, The University, Reading, Berks.
Professor E. L. TRIST, O.B.E., M.A., Graduate School of Business Administration (Management Area), University of California, Los Angeles, California 90024, U.S.A.
Professor A. M. UTTLEY, Ph.D., Laboratory of Experimental Psychology, University of Sussex, Falmer, Brighton, Sussex.
Professor P. H. VENABLES, Ph.D., Department of Psychology, Birkbeck College, Malet Street, London, W.C.1.
Professor M. D. VERNON, M.A., Sc.D., 50 Cressingham Road, Reading, Berks.
Miss M. A. VINCE, M.A., Psychological Laboratory, Downing Street, Cambridge.
Professor D. VOWLES, D.Phil., Department of Psychology, The University, Edinburgh 8.
G. K. WALLACE, Ph.D., Department of Psychology, The University, Reading, Berks.
P. B. WARR, Ph.D., Department of Psychology, The University, Sheffield 10.
Mrs. E. WARRINGTON, Ph.D., Psychological Department, National Hospital, Queen Square, London, W.C.1.
P. C. WASON, Ph.D., Department of Phonetics, University College, Gower Street, London, W.C.1.
A. J. WATSON, M.A., Psychological Laboratory, Downing Street, Cambridge.
R. H. J. WATSON, Ph.D., Department of Nutrition, Queen Elizabeth College, Campden Hill Road, London, W.8.
Miss K. P. WATTS, Psychological Laboratory, Downing Street, Cambridge.
Professor L. WEISKRANTZ, Ph.D., Institute of Experimental Psychology, 1 South Parks Road, Oxford.
Professor A. T. WELFORD, Sc.D., Department of Psychology, University of Adelaide, Adelaide, South Australia.
T. C. D. WHITESIDE, M.B.E., Ph.D., M.B., Ch.B., R.A.F. Institute of Aviation Medicine, Farnborough, Hants.
J. W. WHITEFIELD, M.A., Department of Psychology, University College, Gower Street, London, W.C.1.
P. WHITTLE, Ph.D., Psychological Laboratory, Downing Street, Cambridge.
R. T. WILKINSON, Ph.D., Applied Psychology Research Unit, 15 Chaucer Road, Cambridge.
The Hon. Mrs. M. WILLIAMS, Ph.D., Addenbrooke's Hospital, Lensfield Road, Cambridge.
J. A. WILSON, B.Sc., 32 Cotham Park, Bristol 6.
J. P. WILSON, Ph.D., Department of Communication, The University, Keele, Staffs.
Mrs. M. WYKE, Ph.D., Maida Vale Hospital for Nervous Diseases, Edgware Road, London, W.9.
Professor O. L. ZANGWILL, M.A., Psychological Laboratory, Downing Street, Cambridge.

OCCUPATIONAL PSYCHOLOGY

Acting Editor: ISABEL BLAIN

APRIL and JULY 1968 Volume 42, Nos. 2 and 3

CONTENTS

The Measurement of Occupational Interest—A Symposium.
- Occupational Interests: Retrospect and Prospect. By K. M. MILLER.
- The Brook Reaction as a Test of Interests and Temperament. By A. W. HEIM.
- Work Values and Chosen Careers: Note on a Trial of an American Work Values Inventory with British Subjects. By T. E. CARRUTHERS.
- The APU Occupational Interests Guide: A Progress Report. By M. C. KILLCROSS and W. T. G. BATES.
- The Predictive Value of the Rothwell-Miller Interest Blank. By D. M. NELSON.

Age and Occupations: an Examination of 1961 Census Data. By M. J. SMITH.

Delinquency, Occupation and Temperament. By H. SYLVIA ANTHONY.

Co-Education and Women's Attitudes to Men in Work and Social Life. By R. R. DALE.

Some Characteristics of Business Students in a Technological University. I. Personality Patterns. By A. G. SMITHERS.

A Psychometric Exploration of Some Attitudes to the Managerial Role. By R. A. M. GREGSON and B. D. JAMIESON.

Methods of Scanning a Large Visual Display. By KEITH ELLIS.

Book Reviews.

Annual Subscription 40 *shillings, overseas* 50 *shillings*

NATIONAL INSTITUTE OF INDUSTRIAL PSYCHOLOGY

14, Welbeck Street, London W1M 8DR

PSYCHOLINGUISTICS

EXPERIMENTS IN SPONTANEOUS SPEECH

FRIEDA GOLDMAN EISLER

Psycholinguistics Research Unit, Department of Phonetics, University College London, England

1968, viii + 170 pp., 50s.

This volume presents and discusses experiments aimed at increasing an understanding of the generative processes involved in the production of speech. Special emphasis is placed upon the measurement of extralinguistic phenomena in spontaneous speech but structural aspects have also been studied. The main parameters were hesitation pauses, breathing in speech and sentence complexity. The orientation is towards their interpretation as indicators of cognitive processes and of the higher nervous activity that accompanies the act of speech production. These findings are relevant to all those workers who are interested in evidence about changes in the states of speakers in relation to the properties of the speech and language generated.

CONTENTS
The Problem. Introduction. Some Facts Concerning Pauses in Speech. The Significance of Hesitation Pauses. The Semantic Determination of Pauses and Spontaneity. Complexity of Cognitive Operations, Information and Hesitation. The Cognitive Status of Syntactic Operations. Sequential Temporal Patterns in Speech. Temporal Patterns of Cognitive Activity and Breath Control in Speech. The Significance of Breathing in Speech. Experimental Validation of Speech-breathing Parameters. Conclusion. Appendices. Notes. References. Author Index. Subject Index.

Academic Press **London and New York**

Berkeley Square House, Berkeley Square, London, W1X 6BA, England.
111 Fifth Avenue, New York, N.Y. 10003, U.S.A.

Information Theory of Choice-Reaction Times

D. R. J. LAMING
Psychological Laboratory, University of Cambridge, England

1968,
x + 172 pp.,
63s.

This book represents an important advance towards a comprehensive quantitative theory of choice-reaction times. The central thesis is simple: the response on each trial of a choice-reaction experiment expresses all the information about the signal that the subject has collected. The choice-reaction process should therefore be similar to Stone's model based on the sequential probability ratio test. Certain non-parametric predictions from the theory are confirmed by experiment.
There are review chapters covering Communication Theory and other mathematical models for choice-reaction time, definitive reports of seven new experiments, a detailed mathematical development and appraisal of Information Theory applied to choice-reaction times, and some important and original empirical results on sequential constraints in choice-reaction experiments. An answer is also given to the question: What assumptions about the state of nature are implicit in this application of Information Theory to human behaviour?"

This book is intended primarily for those engaged in research in choice-reaction times and related topics. But this development of Information Theory and the empirical results have a wider importance than this. They will be found relevant by anyone interested in the study of human behaviour in terms of the processing of information.

Chapter Headings

The Explanation of Choice-Reaction Times in Terms of Communication Theory. Alternative Models for Choice-Reaction Times. The Random Walk Model. The Apparatus and the Experimental Procedure. The Experiments. An appraisal of the Random Walk Model. A Critical Examination of Certain Alternative Models for Choice-Reaction Time. Sequential Analysis. Future Developments. Appendix A: The Two-Choice Theory. Appendix B: The Multi-Choice Theory. Appendix C: Two Methods of Analysis. References. Author Index. Subject Index.

Academic Press

London and New York

Berkeley Square House,
Berkeley Square, London, W1X 6BA, England

111 Fifth Avenue,
New York, N.Y. 10003, U.S.A.

Q. Jl exp. Psychol. (1969) 21, 105–126

CONTROL OF MOVEMENT PATTERNS IN ANIMALS†

ROBERT A. HINDE

Sub-department of Animal Behaviour, Madingley, Cambridge

As a more junior Fellow of the same college in Cambridge as Sir Frederic Bartlett, I am especially glad of this opportunity to join with members of the Experimental Psychology Society in honouring him. It would be superfluous for me to speak here of Sir Frederic's academic achievements: so many of our members are engaged in research on problems which sprung from his pioneering work, and his influence on the development of psychology as a science certainly needs no elaboration by me. But I would like to make one personal remark. I suppose all younger men feel that there are just a very few members of the preceding generation who have achieved a stature which will always be beyond the reach of any of their own contemporaries. Sir Frederic has always been one of those for me. I am speaking here not only of his academic achievements, but of his kindness as a man; and I would like to tell him how much I valued his willingness to share some of the benefits of his experience when I was blundering about on the borders of his subject.

Introduction

Work with sub-human species has not played a large part in the discussions of this Society. Most of the problems tackled by its members concern people, and can best be solved with human subjects. With many problems concerning Remembering and Thinking, indeed, they can be solved only with human subjects. But in some fields of research a comparative approach can provide perspective to the human material, and experimental interference of a type not feasible with human subjects is often revealing. I believe that this may be the case with the control of skilled movements. In man, this control involves a number of different principles which it is the psychologist's task to tease apart. Some of these principles are exhibited in a more clear-cut form in animals, and it seems at least possible that the psychologist might be helped by an occasional comparative over-view. In this paper, therefore, I shall attempt to sketch some principles which seem to operate in the control of movement patterns in animals, and in the development of that control. In doing so I shall be drawing very extensively on the work of others, perforce in a rather sketchy fashion: my aim is to emphasize the diversity of principles which must be invoked in understanding the control of movements, and to show how these are being studied in sub-human species. We may consider first movement patterns which are already functional, and then their development in the individual.

---

† The text of the Sir Frederic Bartlett Lecture delivered to the Experimental Psychology Society, 7 January 1969.

Movement Patterns of Adult Animals

It is convenient here to consider successively categories of movement patterns differing in the role of feedback in their control. First, some cases in which patterned output from the central nervous system does not depend on feedback will be described. These will be followed by some cases in which the patterned output is modified by feedback from the periphery, and then by some examples in which external feedback plays the major role.

Central patterning

Even quite complicated movements may be independent of feedback. The swallowing movements of higher mammals are a case in point. Swallowing is quite a complicated movement, involving contractions of about 20 muscles with motor neurones scattered from mesencephalic to third cervical levels. The temporal patterning of the contractions is independent of the means used to elicit swallowing—electrical stimulation of the superior laryngeal nerve with a wide range of frequencies, stimulation of the pharynx with a cotton swab, or rapid injection of water into the mouth. The patterning is thus not imposed by the eliciting stimulus, and arises within the mechanism of swallowing. There is considerable evidence that the patterning is not a consequence of proprioceptive feedback. For instance, if the contraction of one of the muscles is prevented, the pattern of contraction among the rest is not affected. Also a wide variety of treatments likely to upset proprioceptive feedback, including strychnine, cocainization of the pharynx, excision of participating muscles, fixation of the hyoid mass at an extreme position, and traction on the tongue, all fail to produce any major change in the temporal organization of the movement. It thus seems certain that co-ordination is entirely central (Doty and Bosma, 1956; Doty, Richmond and Storey, 1967).

Another example is provided by the vocalizations of adult birds. Experiments with a number of species have shown that, once they are fully developed, the vocalizations of many adult birds are affected little if at all by deafening. For instance if adult cocks are deafened, they still crow and produce other calls in the appropriate functional context. Similar results have been obtained with the chaffinch (*Fringilla coelebs*), which has quite a complicated song. Once the song is fully developed, its performance (including both the characteristics of the song itself and the temporal pattern of singing) is independent of auditory feedback. Of course other sources of feedback could be responsible for the patterning; but severance of the hypoglossal nerve which innervates the syrinx, although affecting the syllable structure and tonality of the song, does not affect the pattern of phrasing, duration or stability of the songs (Konishi, 1963; Nottebohm, 1967, 1968; Konishi and Nottebohm, 1969). While it remains possible that the characteristics of the singing would be affected by *distorted* feedback, it seems clear that feedback is not essential for the maintenance of most of the characteristics of normal patterning of the output.

Interaction between central and peripheral factors

While the form of some movements is thus not affected by the removal of at least the more important sources of feedback, many movements depend on a complex interplay between central patterning and proprioceptive feedback. We may consider some examples.

The flying movements of locusts involve alternating up and down movements of the wings. It was earlier thought that the alternation of contractions in the elevator and depressor muscles was due to proprioceptive feedback loops, but it is now clear that this is not the case. Although the stretch receptor in the hinge of the wing does provide a sensory discharge, its removal results only in a slowing of the discharge from the thoracic ganglia, and not in a change in its patterning. Indeed in a preparation which consists in essence only of the head and ventral cord, a non-phasic stimulus of wind on the head results in a phasic output from the thoracic ganglia similar to, though slower than, that occurring in normal flight. Detailed analysis indicates that the flying movements depend on a number of central pacemakers, coupled into a dominant frequency which depends on the quantity of the input. These may involve excitatory couplings between synergistic groups of motor neurones, and inhibitory interactions between the antagonistic groups. The advantages of this system over a chain reflex may lie in the possibilities of adjusting wing-beat frequency: at high frequencies the downstroke muscles start to contract before the end of the upstroke, and thus help to decelerate it (Wilson, 1961, 1964; Wilson and Wyman, 1965; Weis-Fogh, 1964).

Although the sensory discharge from the hinge affects the frequency but not the patterning of the thoracic discharge, other sensory stimuli can have a profound effect. In particular optic input, by which the locust usually maintains its equilibrium, affects the relative discharge rate on the two sides in such a manner that the sensory input is balanced [Fig. 1(a), (b)]. Wilson (1968) therefore describes the motor output as depending on an interaction between a centrally determined "motor score" and a mechanism correcting deviations from a "sensory template."

In the swimming of the dogfish, phasic feedback plays a somewhat more complicated role. In normal swimming, waves of contraction in the body muscles pass down the length of the body: there is a phase difference between the two sides of the body, giving the sinusoidal swimming pattern. Lissmann (1946*a*, *b*), studying the role of deafferentation, found that the swimming rhythm continued in partially deafferentated spinal fish, provided that a few dorsal roots remained intact. He concluded that sensory feedback was necessary for the maintenance of the rhythm. But, more recently, Roberts (1969) found that the rhythmical output from the spinal motor neurones persisted even in curarized fish, in which sensory feedback was presumably absent. However the rhythm was slower and less persistent than in normal fish. Furthermore the motor units did not fire successively down each side of the body, as in normal swimming, but simultaneously. It thus seems that the swimming pattern depends on an interaction between an autonomous central output and feedback from the contracting muscles: in the normal fish the phasic input from swimming overrides the spontaneous output of the motor neurones, entraining them to fire at the right frequency and organizing the longitudinal propagation of contractions down the body.

A comparable interaction governs the locomotor rhythms of anurans. Székely (1968) implanted segments of brachial spinal cord into the dorsal fins of salamander larvae, sealing them in place with limb grafts. The limbs subsequently

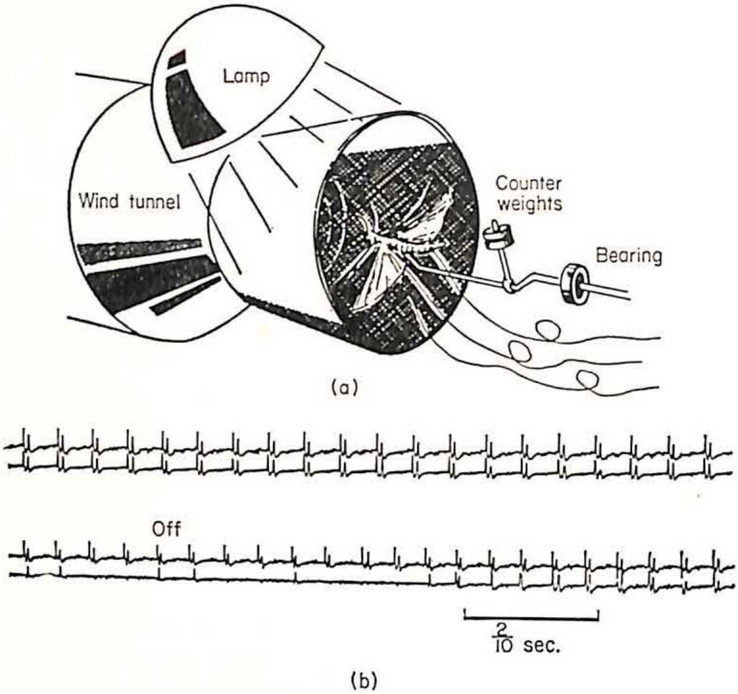


FIGURE 1. (a) Apparatus for observations on the rolling behaviour of locusts. The lamp and horizon-simulating section can be rotated as a unit. After the animal is mounted, the balancing weight is adjusted so that there is no roll torque in any position when the animal is quiet. When the animal rolls the electrode wires twist about the roll bar, but four to five turns are possible before the wires limit the animal's ability to roll. (Reproduced with permission from Wilson, 1968.) (b) Discharges to the two hind subalar muscles of a locust with one hind wing removed. In the dark (upper records) the discharges to the two sides are similar. When the light is switched on (beginning of lower trace), the discharge becomes asymmetrical in such a way that the muscle to the intact wing is less active than usual. Turning the light out ("Off") results in a return to symmetrical output. (Reproduced with permission from Wilson, 1968.)

showed spontaneous rhythmic movements when the cord segment received only non-phasic input. Thus co-ordinated movements of single limbs could be controlled by brachial segments of spinal cord receiving only non-rhythmic input. The inter-limb co-ordination involved in the normal diagonal pattern of locomotion, however, seems to require rhythmic input. By means of deafferentation experiments Gray and Lissmann (1946*a*, *b*) showed that it requires the presence of at least one intact spinal nerve. Their experiments indicated that the rhythmicity of the input, rather than the total amount, was the crucial issue. Thus here again the patterning depends on an interaction between central and peripheral factors.

Such examples indicate that the interactions between centrally patterned output and peripheral feedback may occur in a number of ways In the case of rhythmically recurring responses some of the possibilities are illustrated by Bullock (1961) in Figure 2. In the first case the timing cues come from the periphery, though the patterning of the response is determined centrally. In the second, the response is initially triggered from the periphery, but then recurs as a consequence of feedback. In the remaining cases the timing cues come from a

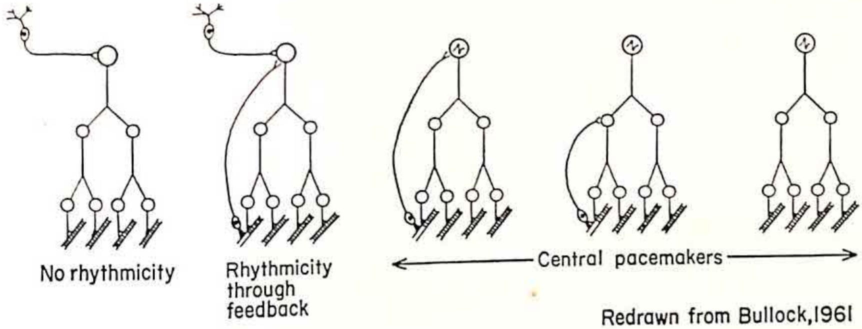


FIGURE 2. Diagrams of hypothetical types of nervous mechanisms producing patterned movement. (Reproduced with permission from Bullock, 1961.)

The three levels of neurones represent branching chains in whose junctions integrative properties may alter the actual impulses and deliver them to the effectors. The first two are shown with exteroceptors, the others with spontaneous pacemakers. The second and third have proprioceptive feedback acting on the trigger neurone, while in the fourth it influences only the patterning of the response, but not its timing.

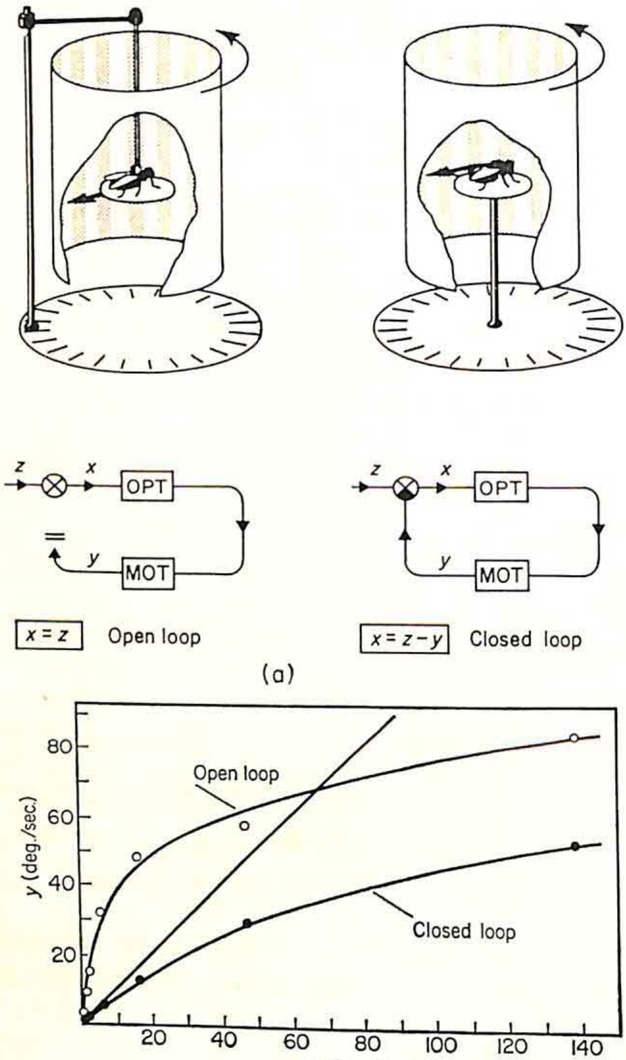
central pacemaker. Feedback from the periphery may affect either the pacemaker itself, thereby modifying the temporal characteristics of the rhythm; or one of its followers, thereby modifying the form but not the frequency of the rhythmic output. In some cases feedback may be absent. It goes without saying that Figure 2 represents only hypothetical simplified systems: in practice feedback effects may be positive or negative, come from varied sources, and operate at a number of different points in the system.

Oriented movements

So far we have been concerned with movements lacking specific orientation to the external world. When a movement is oriented in space, continuous feedback is usually involved, though orientation sometimes depends only on ballistic control: for instance the orientation of a firefly towards a flash which is over before the orientation movement starts cannot depend on feed-back control. Again, the extension of the forelegs of the praying mantis takes place so rapidly that it must depend on an open system (Mittelstaedt, 1962).

Most oriented movements, however, are controlled by feedback. A simple and classic case is the optomotor response of a fly placed in the centre of a cylinder painted with vertical stripes [Fig. 3(a), (b)]. Under normal conditions, the fly's angular velocity is approximately equal to that of the drum at low drum speeds. If the fly's body is held still and its walking movements allowed to rotate a turntable underneath it, the loop is opened. Under these conditions, the fly turns much faster than does the drum (Mittelstaedt, 1964).

Now when an insect walks, there is a continuous displacement of the image of its surroundings across the surface of its receptors. We must therefore ask, why does this relative movement not induce forced turning movements, as in the optomotor response? The absence of such an effect is not due to a mere blocking of the optomotor reflex, for the following reason. If the head is rotated through 180° about the axis of the neck, so that the positions of the eyes are reversed, the



(a)

(b)

FIGURE 3. (a) Measurement of the optomotor response of a fly under "open loop" (upper left) and "closed loop" (upper right) conditions. In the former the fly is fastened by its thorax and the rotatory velocity of a paper disc held by the fly is measured. In the latter the fly moves freely on a standing platform and its velocity relative to the ground is measured. In the diagrams of the control patterns, OPT = afferent causal relation; MOT = efferent causal relation.
(b) The results of the two experiments. The rotatory velocity (y) of the fly is plotted against that of the cylinder (z). (Reproduced with permission from Mittelstaedt, 1964.)

fly now moves in the opposite direction to the cylinder. However if such a fly starts to move of its own accord while the cylinder is stationary, it spins rapidly in small circles until exhausted. On the basis of this and other experiments, von Holst and Mittelstaedt (1950; and von Holst, 1954) suggested that such orientational movements are the result of a comparison between the actual afferent stimulation and that which should have occurred according to the momentary motivational situation. The control system must therefore be of the type shown in Figure 4. The order to the effectors depends on the difference between a

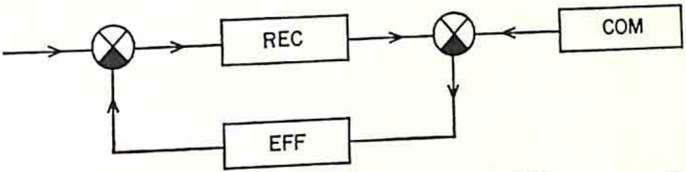


FIGURE 4. General characters of a control system with variable command. The input to the effectors (EFF) depends on the difference between signals from the receptors (REC) and the command (COM).

correlate of the movement of the image on the receptors and a correlate of the command—in the case of the optomotor response the command is zero, while when the head is reversed the sign of the input from the sense organ is reversed so that in a self-initiated movement the effectors receive the sum of two inputs rather than their difference.

The perceptual side of such a response has been studied in the beetle *Chlorophanus* by Hassenstein (1959, 1961; Hassenstein and Reichardt, 1959), who used a wide variety of stripe patterns to elicit it. In order to account for the data they proposed a model in which signals from neighbouring ommatidia are transmitted through parallel filters with short and long resonance times, and then via multiplication units to the effectors (Fig. 5). Response thus depends on the signal

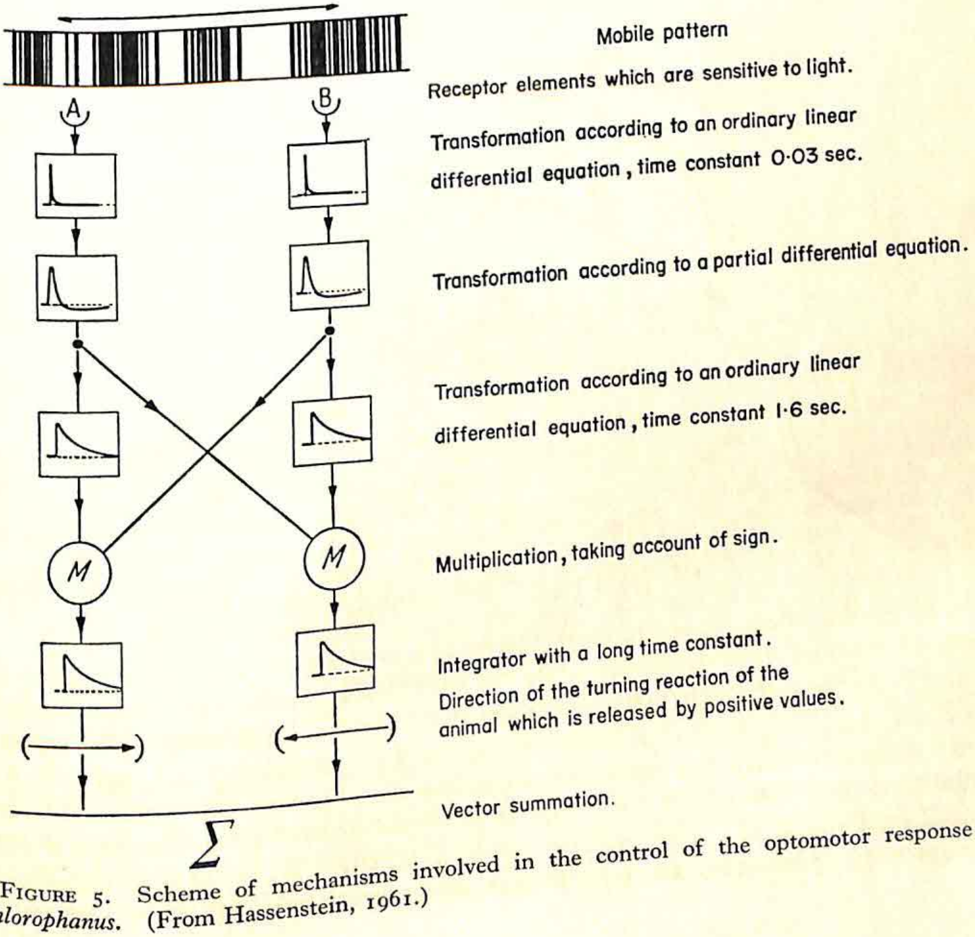

Mobile pattern
A
B
Receptor elements which are sensitive to light.
Transformation according to an ordinary linear differential equation, time constant 0·03 sec.
Transformation according to a partial differential equation.
Transformation according to an ordinary linear differential equation, time constant 1·6 sec.
M
M
Multiplication, taking account of sign.
Integrator with a long time constant.
Direction of the turning reaction of the animal which is released by positive values.
Vector summation.
Σ

FIGURE 5. Scheme of mechanisms involved in the control of the optomotor response in *Chlorophanus*. (From Hassenstein, 1961.)

from one ommatidium reaching the multiplication unit before the signal from a nearby previously stimulated ommatidium has decreased to zero, giving a non-zero product.

Returning to the motor side, more complex problems arise when we consider not the orientation of the body as a whole, but that of mobile appendages. We may consider just one example. The strike of the praying mantis depends overall, as we have seen, on an open system. However the head is mobile on the thorax, and thus the deviation of the strike from the thorax, which carries the legs, must be controlled with respect to the optic axis. Mittelstaedt considered three possible control mechanisms (Fig. 6).

(1) The head position is measured proprioceptively and added to the optic information, so that the forelegs strike appropriately.

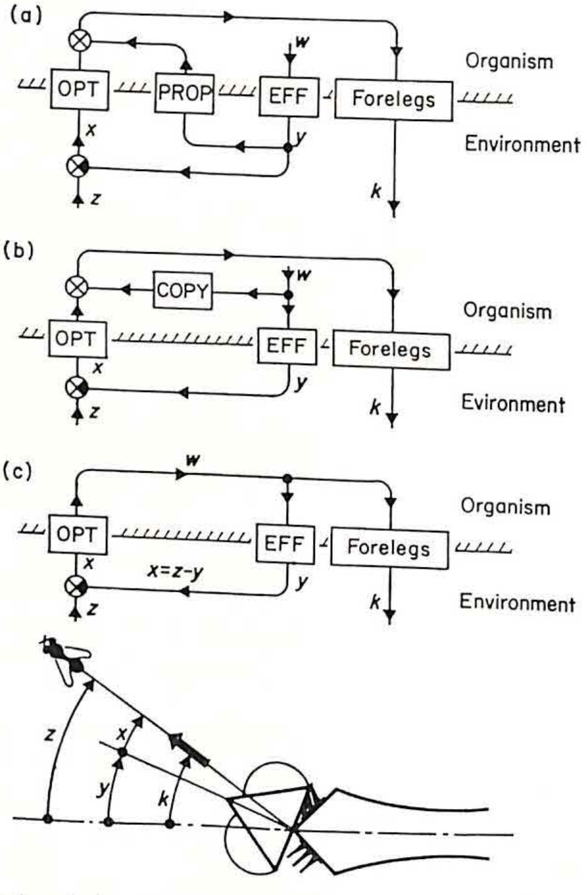


FIGURE 6. Three basic solutions for the problem of prey localization in *Mantis*. (Reproduced with permission from Mittelstaedt, 1962.)

OPT: sub-system which yields information about deviation x of the prey from the eye axis; PROP: sub-system which yields information about deviation y of head from body axis; EFF: sub-system which controls deviation y of head according to its input variable w, the "order"; COPY: sub-system which transmits a "copy" of the order w to the optic output; FORELEGS: sub-system which controls the deviation k of strike from the body axis; z, deviation of prey from body axis. (a) Proprioceptive mesh, (b) copy-mesh, (c) loop with branch-off.

This possibility is ruled out because, if the head is fixed 10 to 20° to one side, the animal misses the prey to the opposite side. It thus behaves as if it did not notice the deviation.

(2) A copy of the input of the system controlling head orientation is added to the optic output.

(3) The optic output is fed back negatively into the system controlling head orientation so that the head turns towards the prey. The final steady state of this optic loop then determines the direction of strike.

Since the mantid does fixate its prey before striking, the last hypothesis is the most likely. However this can not be the whole story, for unilateral deafferentation of the neck proprioceptors results in an error in striking to the opposite side: the proprioceptors must therefore play some role. Such animals show a permanent neck deviation, and this is presumably responsible for the error. Thus neck muscles and proprioceptors form a loop within the optic loop (Fig. 7). On this view the contraction of the neck muscles is adjusted according to the information provided by the optic loop. Movement ceases when this order is balanced by information from proprioceptors indicating the head position. The same optic output is then used to determine the deviation of the forelegs from the body's axis (Mittelstaedt, 1960, 1962).

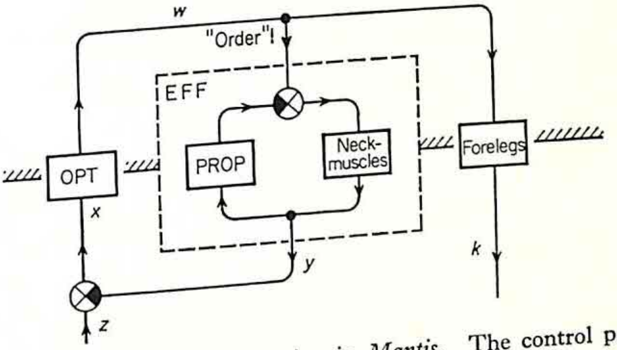


FIGURE 7. Control pattern of prey localization in *Mantis*. The control pattern of EFF (see Fig. 6) is a loop which turns the head into the position determined by the order *w*. (Reproduced with permission from Mittelstaedt, 1962.)

The few studies just cited illustrate the sort of work which is in progress in this field. The work of Horridge and his collaborators on the control of eye-movements in the crab *Carcinus* provides elegant examples of the progress possible with this sort of approach (Horridge, 1966; Horridge and Burrows, 1968; Horridge and Sandeman, 1964). That similar methods are fertile also with sub-human vertebrates is illustrated by the work of Mittelstaedt (1964) and by studies cited later in this paper.

Thus with certain exceptions, feedback plays an essential role even in these rather crude movements of lower animals. In higher animals, although the elicitation (Gorska and Jankowska, 1961) and even the acquisition (Taub, Bacon and Berman, 1965; Taub, Ellman and Berman, 1966) of simple movements can occur when the limb is deafferentated, feedback usually plays a vital role. Indeed the skilled movements of human subjects probably depend on series of hierarchically organized closed loop systems (e.g. Welford, 1968), with continuous monitoring of the state of muscular contraction (Matthews, 1964) playing an important role.

Development of Movement Patterns

That even species-characteristic movements may have a long developmental history is shown, for instance, by Bernstein's (1967) careful descriptions of the changes in walking with age. Such changes suggest that a number of different processes may be involved in the development of the adult pattern.

Development independent of "practice"

In the first place, in lower vertebrates the precision of some quite complicated and finely adjusted movements may be independent of experience. Though now classic, Sperry's (eg. 1950, 1951) experiments on visuomotor co-ordination in amphibians must be mentioned here. Some were concerned with the way in which the retina becomes connected in an orderly manner with the visual regions of the brain. To assess this, Sperry used both the optomotor and prey-catching responses. If the eyeball is rotated in its orbit, these responses are, as would be expected, reversed. If the optic nerve is severed and then allowed to regenerate to the rotated eyeball, the responses are still reversed. Thus although the axons are tangled in an apparently random fashion in the scar region, the anatomically correct connections must have been re-established, even though, owing to the rotation of the orbit, they are functionally incorrect. On the basis of this and other experiments, Sperry argues that the axons must be functionally differentiated from each other, and that the orderly patterning of synaptic connections in the brain is determined by biochemical affinities specific for each part of the retina. Although the precise mechanisms here are still not clear (see e.g. Wolstenholme and O'Connor, 1968), the general principle that the development of anatomically correct nervous connections may be independent of experience is well established. This must apply not only to the lower vertebrates but also to higher mammals and to man.

Role of reafference

We saw earlier that even relatively simple organisms behave differently to relative movement of the environment caused by movement of their own bodies from that caused by movement of the environment itself. To account for this, it has been suggested that the change in stimulation received is compared with a representation of the required change in stimulation (which may, of course, be zero), further action depending on the results of the comparison. This was discussed earlier in the context of von Holst and Mittelstaedt's analysis of the optomotor responses of insects. Turning to a similar problem, Figure 8(a) shows the control system suggested to account for the relation between eye-movement and perception in man. Any command to the eye muscles is accompanied by a corollary discharge which is compared with the change in stimulation received as a result of movement. A change in sensory input produced by an

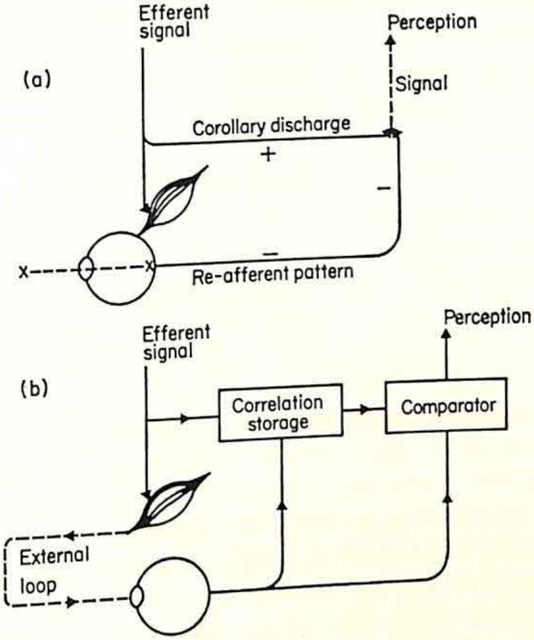


FIGURE 8. (a) Basic relations between eye-movements and perception. (Redrawn with permission from Teuber, 1961.) (b) Relations between eye-movements and perception required to account for consequences of experimental rearrangement, disarrangement and deprivation. (Reproduced with permission from Hein and Held, 1962.)

external change is not accompanied by a corollary discharge and so is registered in perception as due to movement in the external world. If, however, the change in stimulus input is a result of voluntary movement of the eyeball, its consequences are cancelled by the corollary discharge, and no movement of the external world is perceived.

But this poses a developmental problem. How does the corollary discharge from the efferent system resulting from a command to, say, move the eyes 10° to the right, and the reafference from a shift of 10° of the image on the retina, come exactly to balance each other? Is the coding of the messages determined by the growth processes of the organism, or is it a result of experience? Held, Hein and their collaborators, following up the work of earlier investigators, have shown that in man and higher mammals such correlations are established as a result of experience during self-initiated movements. If the visuomotor relations are re-arranged, as for instance when the subject wears prismatic spectacles, readjustment is possible if the subject is allowed to move about, but not if similar movements are induced passively. The readjustment depends on comparison between intended movements and the reafferent stimulation: since the prisms produce a regular displacement of the image, this can be "allowed for" and visuomotor co-ordination restored. Hein and Held (e.g. Held, 1961) therefore suggest the addition, to the von Holst/Mittelstaedt model, of a "correlation storage unit" where traces of previous combinations of concurrent efferent and reafferent signals are retained [Fig. (8b)]. The correlated discharge of a command to the efferent system is supposed to select from the store that combination containing the identical efferent part and to activate the corresponding reafferent trace. This is sent to the comparator for comparison with the current reafferent signal. If the sensory field is distorted, new combinations will be stored, and thus more than one combination will be available in the store. Since readjustment occurs, selection is presumably weighted by how recently the combinations were stored. With constant rearrangement, selection of old combinations gradually ceases. If in the new conditions the discrepancy between new and old reafference is not predictable, as for instance with a prism which changes continuously in power, no rearrangement is possible.

Held and Hein (1963) have shown that similar principles apply also to the early development of visuomotor co-ordination in ontogeny. Kittens were reared in the dark and then given visual experience, in pairs, in an apparatus which allowed one kitten to move about actively, while the other was confined in a "gondola" and moved only passively as a consequence of the movements of its active partner. Thus both kittens received approximately the same amount of visual stimulation, but in one case the stimulus changes were a consequence of the animal's own locomotion, while in the other they were passively induced. When subsequently tested on a visual cliff, or for extension of the paws as they were lowered passively towards a solid surface, the active kittens were uniformly superior. Similarly, differences could be induced between two eyes of the same kitten if one eye was permitted stimulation only during active locomotion and the other during passive. Furthermore kittens raised without sight of their forelimbs, but able to engage in self-initiated locomotion, did extend their paws when lowered towards a horizontal surface, but were unable to guide their paws towards the solid parts of an interrupted surface [Fig. 9(a), (b)]. Thus the ability to orient the paw towards a visible target requires an integration of sensori-motor systems not necessary for the extension response, and apparently requiring vision of the paw during self-initiated movement (Hein and Held, 1967). Similar results have been obtained by Held and Bauer (1967) for monkeys reared under

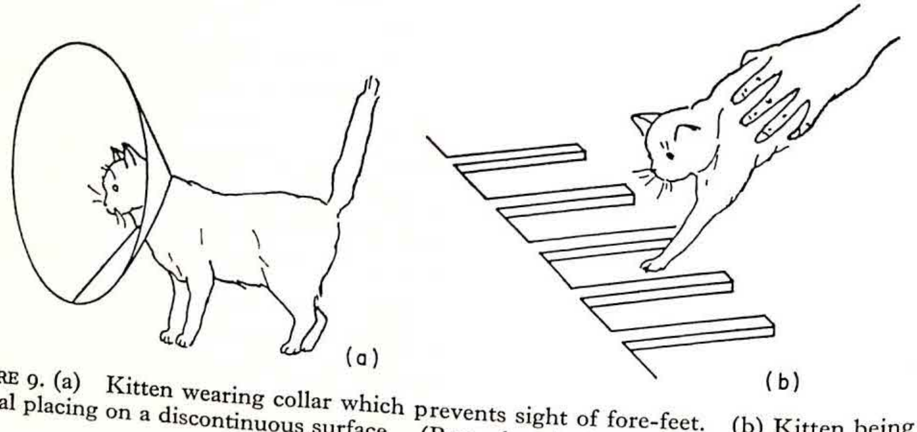


FIGURE 9. (a) Kitten wearing collar which prevents sight of fore-feet. (b) Kitten being tested for visual placing on a discontinuous surface. (Reproduced with permission from Hein and Held, 1967.)

conditions in which they could not see their arms. Such data suggest that movements oriented with respect to the environment depend on a number of interacting systems, developing with some degree of independence, and perhaps organized hierarchically.

That similar effects occur in lower vertebrates is suggested by experiments involving orientation in space after unilateral blinding or labyrinthectomy in amphibians. For instance Jahn (1960) has shown that certain frogs regain a symmetrical posture in about 10 days after unilateral blinding and about 3 months after unilateral labyrinthectomy. That the compensation involves the re-setting of a control system is indicated by the finding that temporary blinding is followed by asymmetry in the opposite direction when sight is restored. That it depends on input through the intact organ is shown by the effect of the light period on recovery: after unilateral blinding normal compensation occurs with 30 min. of light per day, but with shorter periods of light, recovery is retarded.

Role of conditioning

So far, in considering the development of movement patterns, we have seen that although in some cases precise orientation with respect to the environment can be determined almost entirely by maturational processes within the central nervous system, it usually depends also on experience during active movements. We shall now consider two further cases which illustrate additional ways in which experience may affect movement patterns.

The first concerns the licking behaviour of rats and illustrates the manner in which classical and operant principles may be interwoven in the modification of response patterning. When given a water-spout, rats lick in rhythmic bursts. The response of licking is elicited by a fairly specific stimulus situation, and each drop of fluid elicits a rhythmic burst of licks whose number increases with the size of the drop. These facts, together with the finding that licking can be classically conditioned, suggests that licking can be regarded as a respondent. By contrast the overall licking rate (including both the number of bursts and the pauses between them) is affected by deprivation, the sucrose concentration, and the schedule of reinforcement. Thus this characteristic of drinking behaviour, as well as the initial approach to the drinking tube, behaves like a conventional operant response. Although some anomalies remain (e.g. overall licking rate decreases with increase in the size of an FR schedule), it is clear that some insight into the control of licking can be gained by regarding it as involving priciples of both classical and operant conditioning (Hulse, 1967).

This analysis emphasizes that the distinction between classical and operant conditioning depends in part of the selection of response measures which conform with one paradigm or the other: in the development of naturally occurring movement patterns both are closely interwoven. Whereas in most laboratory studies of operant conditioning there is no initial relation between cue stimulus and operant response, in nature there is often an initial relationship between an eliciting stimulus and the response (e.g. Thorpe, 1963). For instance, young nidifugous birds peck at a wide variety of objects which contrast with their background: although pecking may not lead immediately to the ingestion of edible objects, it continues for a while. This is functionally important because it permits discrimination between edible and inedible objects to develop. Similar principles may apply at many stages in the development of complex movements.

The second case of the role of learning in the development of a species-characteristic movement pattern concerns song-learning in birds. Singing is, of course, a form of motor skill, even though it does not involve orientation with respect to the environment, and it is one in which the form of the movement can be accurately traced by the use of sound spectrographic techniques. The most interesting data concern song-birds whose species-song is dependent on learning, but which will not learn to imitate any song that they hear. One of the best studied species is the chaffinch (*Fringilla coelebs*), and the initial facts are as follows (Thorpe, 1961; Fig. 10).

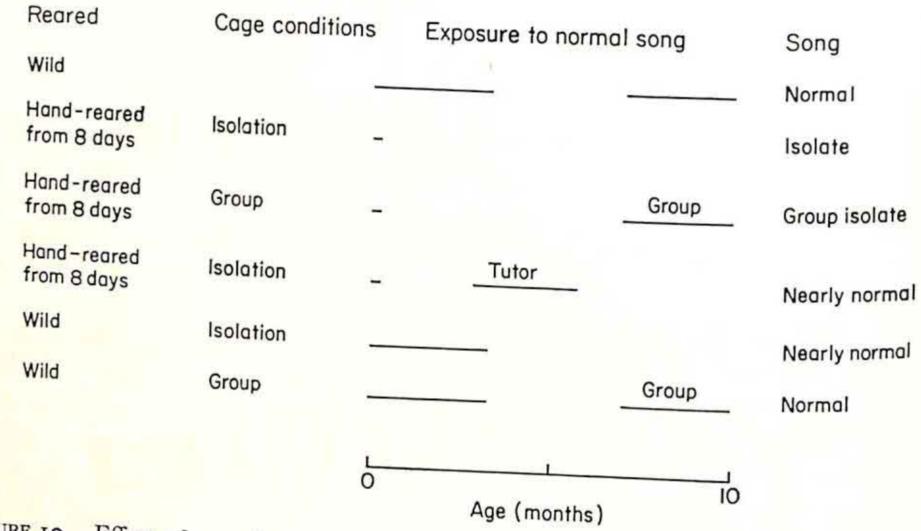


FIGURE 10. Effect of experience on song development in the chaffinch. (Data from Thorpe, 1961). Line indicates exposure to song in natural conditions (unlabelled), song from Group companions, or song Tutor.

(i) Each adult male normally has a repertoire of several songs of the type shown in Figure 11(a). These develop from a rambling subsong of indefinite length by progressive omission of extreme frequencies and approximation towards the normal pattern. This usually occurs under the influence of testosterone from the developing gonads when the birds are about 10 to 12 months old.

(ii) Individuals hand-reared in isolation from other individuals develop only a simple form of the song which is not divided into phrases [Fig. 11(b)].

(iii) Chaffinches hand-reared in groups, but isolated from normal song, develop songs more elaborate than those of individuals reared in isolation. The songs are similar within groups, but differ between groups: this indicates that learning occurs within each group.

(iv) Individuals trapped in their first autumn and kept in isolation thereafter develop songs more or less within the normal range. The difference between

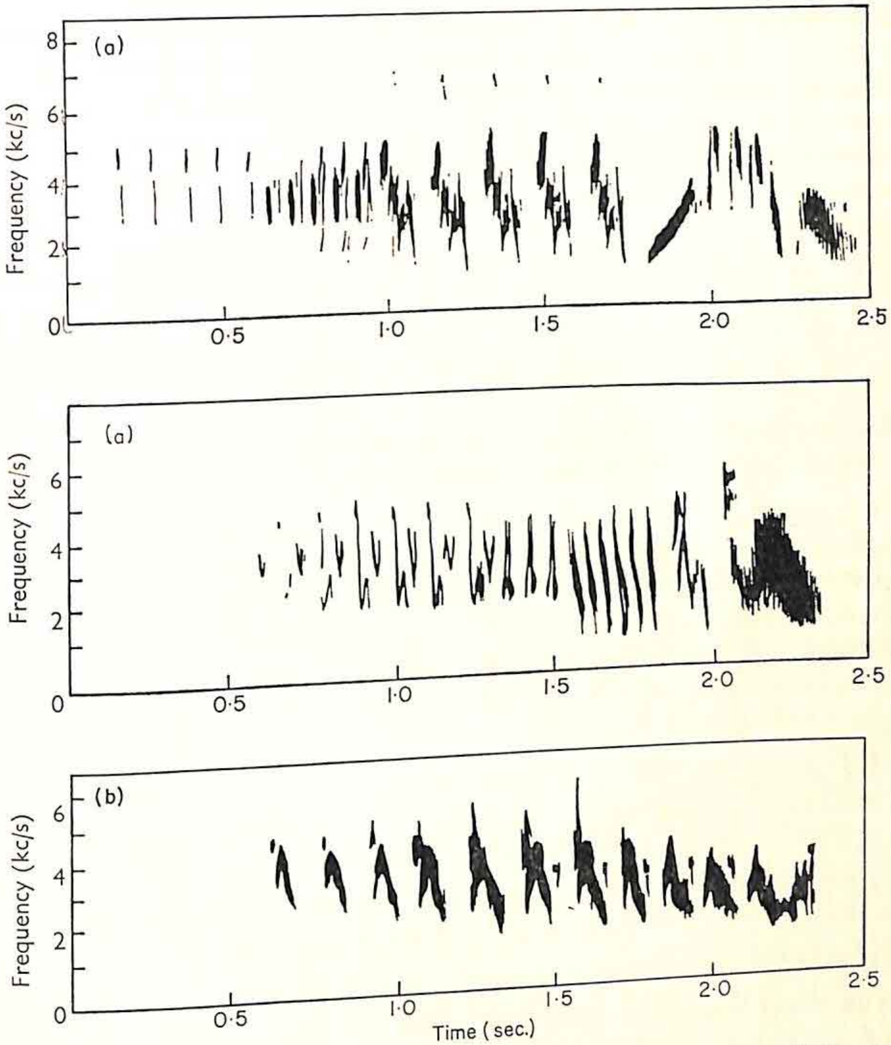


FIGURE 11. Sound spectrograms of chaffinch song. (a) Wild-type songs. (b) Song of bird hand-reared from 6 to 8 days and kept subsequently in auditory isolation. (Reproduced with permission from Thorpe, 1961).

these birds and those hand-reared from an early age [(ii) above] indicates an effect of experience during the first few months of life, long before the birds themselves begin to sing.

(v) Chaffinches caught in their first autumn and kept subsequently in groups develop songs just like those of wild birds. The difference between these birds and the autumn-caught isolates [(iv) above] indicates the importance of experience during the period of song development: this probably occurs in the course of countersinging.

(vi) Hand-reared isolates, exposed to tape recordings of chaffinch song during their first autumn and winter, subsequently produce near-normal songs. Such birds will not learn anything which is played to them. A chaffinch song played backwards, or with the end in the middle, is imitated, but an artificial chaffinch song with notes of pure tonal quality is not. The wider range of sounds in the subsong shows that the restriction in what they will learn is not a matter of the range of sounds which the bird's own syrinx can produce: Thorpe suggested that the criterion may be a resemblance in note-structure to normal chaffinch song.

(vii) Chaffinches isolated only from their first autumn are more restricted in what they learn when tutored. They will not learn reversed or re-articulated song, but are affected by hearing normal chaffinch song. The greater selectivity of these birds, as compared with the hand-reared isolates, must be due to experience before they themselves started to sing.

(viii) Once a chaffinch has acquired its repertoire of songs in its first breeding season, little further change takes place.

These results indicate that song development depends on experience at two periods of the life history—early on, before the bird itself has started to sing; and in the first spring, in the course of countersinging with other individuals. One possibility is that the effect of early experience is to increase the reinforcing effectiveness of song. On this view, the development of full song from subsong would be due to the reinforcing value to the bird of hearing itself sing the full song. Two predictions would follow from this:

(1) deafening the bird after it had experienced song, but before it started to sing itself, would make it impossible for the bird to monitor its own output and thus would nullify the effects of the early experience;

(2) since reinforcers are usually effective for a wide variety of responses, any reinforcing effects of hearing song on singing should also operate for other responses.

Both of these predictions have been confirmed. We may consider first the effects of deafening (Konishi and Nottebohm, 1969). Chaffinches which have already developed their songs before deafening maintain them subsequently, relatively unaltered. The effects of deafening before song has fully developed depend on the time of the operation: the earlier the stage of song development, the greater its effects. Individuals deafened at about 100 days of age, who had heard no chaffinch calls other than their own nestling, fledgling and perhaps juvenile calls, subsequently produced only very unstructured and rudimentary songs. Birds deafened in their first winter, with some previous experience of subsong, produced songs with syllables more like those of wild birds, but of considerable simplicity. Birds deafened during full song development produced songs much nearer the wild type; but even if the operation was performed after several days of practising the full song—though before it had reached its final stereotyped form—the song subsequently underwent considerable deterioration which was not reversed by testosterone therapy (Nottebohm, 1967).

These experiments thus showed that the effects of deafening depend on the stage of vocal development at which the operation occurs: there was also strong evidence that sound patterns developed before the loss of hearing may be incorporated, perhaps in a modified form, into the full song. Song development is thus a gradual integrative process, intimately bound up with the development of other vocalizations (Nottebohm, 1967).

The second prediction was that song should serve as a reinforcer for other responses. Stevenson (1967, 1969) tested this with chaffinches kept in cages containing four transverse perches. The two end-perches operated micro-switches, which in turn operated counters: in addition, that end-perch which had been the less preferred during preceding preliminary observations could operate a tape-loop which played a single chaffinch song. Distribution of perching between the two end-perches during stimulus conditions in which every fifth or sixth response to one of the end-perches did produce song, was compared with that during control conditions when perching was simply recorded. A reinforcing effect of song would be indicated by an increase in the proportion of responses to the active perch during the stimulus conditions.

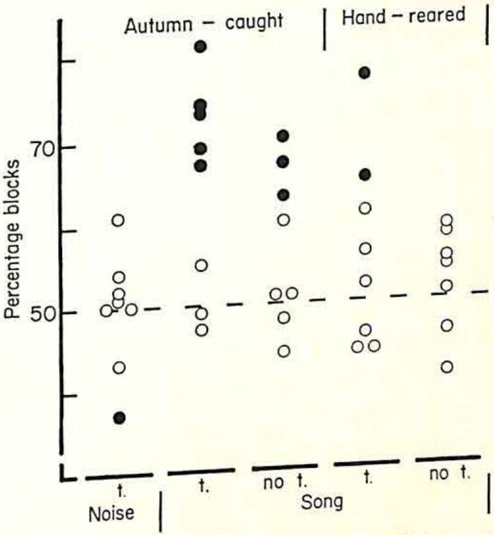


FIGURE 12. A measure of the reinforcing effect of white noise or normal adult song for first-year male chaffinches. Each point represents a different individual, and filled circles indicate a significant difference from chance performance (sign-test, $P<0{\cdot}05$, one-tailed). Birds were either autumn-caught or hand-reared, and injected or not injected with testosterone. (Reproduced with permission from Stevenson, 1969.)

As shown in Figure 12, song was found to be a reinforcer for autumn-caught chaffinches treated with testosterone. A burst of white noise of similar duration and intensity did not show reinforcing properties. Furthermore, none of the eight chaffinches hand-reared in isolation from chaffinch song and tested without testosterone treatment behaved as though song were a reinforcer. Both hand-reared testosterone-treated, and autumn-caught non-injected, groups showed intermediate results. Thus two of the conditions upon which normal song development depends—experience of song and testosterone—also appear to affect the reinforcing effectiveness of song.

One way of relating all these findings is to suppose that song development involves a "model" or "template" which is influenced by experience—particularly experience of the species-characteristic vocalizations. Sounds which approximate to this model are reinforcing and modify the template. Of course the postulation of a model is merely a device for tying together the experimental data: the corresponding physiolgocial reality is quite another issue. However it also permits us to relate song development to certain data concerning the elicitation of song in the adult: that song-type which is most frequent in each individual's repertoire is also the most effective in eliciting song from that individual. This suggests that sound patterns which approximate to the model are both reinforcing and effective in eliciting song (Hinde, 1958).

Studies of other passerine species have shown that song-learning occurs in a basically similar way in a number of other species, though in each the process has its own peculiarities (see e.g. reviews by Konishi and Nottebohm, 1969; Immelmann 1969). For instance, while in the chaffinch the restriction on the sounds which will be imitated and incorporated into the full song seems to depend on their resemblance to normal chaffinch sounds, in other species there is a strong tendency to imitate the song the father sings. Thus bullfinches (*Pyrrhula pyrrhula*) reared by canaries develop a canary-like song even if bullfinches are singing in the same room (Nicolai, 1959; see also Immelmann, 1969).

Species differences are found also in the nature of the sensitive period for song learning. In the chaffinch, song learning comes to an end when full song develops in the first breeding season. A chaffinch, castrated before coming into song, retained the ability to learn a normal song pattern when it was brought into song at two years of age by testosterone therapy. Thus the end of the sensitive period is not age-dependent, but depends on song development itself. Perhaps hearing its own song produces a final adjustment to the template. In other species the sensitive period for song learning ends before the onset of full song. For instance the white-crowned sparrow (*Zonotrichia leucophrys*) has rather distinctive local dialects. Birds taken from the nest and reared in the laboratory develop song patterns which lack the characteristics of the local dialect (Fig. 13). If they are brought in a few weeks later, their song develops normally and they develop the characteristics of the dialect of their own locality. In tutoring experiments, birds exposed to recordings of a particular dialect during the first weeks of life subsequently reproduce it, but birds brought into the laboratory after about three months of age cannot be taught a fresh dialect (Marler, 1967). Since the sensitive period during which white-crowned sparrows can learn a fresh dialect ends before they themselves start to sing, it cannot be ascribed to motor fixation, as in the chaffinch. Konishi and Nottebohm suggest that there are two sensitive periods, one for acquisition of the template and one for the motor learning: in the chaffinch the two end at the same time, but in other species they do not. It is of particular interest that, in the white-crowned sparrow, deafening after the sensitive period but before the onset of singing annuls the effect of the earlier experience (Konishi and Nottebohm, 1969).

The role of group-rearing also raises some interesting problems. Grouping of hand-raised chaffinches leads to the production of more elaborate song patterns than are produced by isolated individuals. No such effect occurs in white-crowned

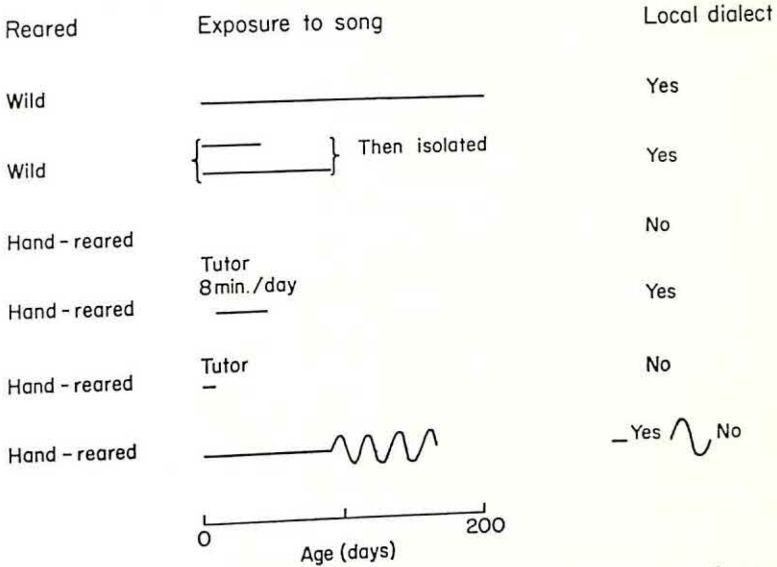


FIGURE 13. Effect of experience on development of song of white-crowned sparrow. (Data from Marler, 1967.) Straight and wavy lines indicate different dialects.

sparrows (Marler and Tamura, 1964), but group-reared Arizona juncos (*Junco phaenotus*) produce more complex syllables and larger repertoires than isolates. Such data suggest that stimulation received in the course of counter-singing in the group situation leads to greater "improvization." However comparative evidence emphasizes the need for caution here: the Oregon junco (*Junco oreganus*), which in nature has a simpler song than the Arizona junco, produces more complex songs if reared in isolation than it does in nature. Individuals reared in groups are more likely to produce near-normal songs than are isolates: mutual stimulation thus serves to bring the course of development back to a normal pathway, rather than to produce further elaboration (Marler, 1967).

Song-learning is also of interest in connection with the role of performance in learning. We have seen that, in both chaffinches and white-crowned sparrows, some learning takes place when the bird is quite young, long before it starts to sing. In that this learning does not involve the response by which its occurrence is subsequently revealed, parallels can be found in studies of perceptual learning (e.g. Gibson, Walk and Tighe, 1959), discrimination learning (e.g. Sutherland, 1959; Mackintosh, 1964) and latent learning (e.g. MacCorquodale and Meehl, 1954; Thistlethwaite, 1951.) It will be apparent that bird song involves a skill which, though lacking orientation to the environment, resembles many human skills in that its acquisition involves learning both the nature of the task and how to carry it out, and that, although the two processes interact, considerable separation is possible.

In conclusion, there is of course no suggestion of a close parallel between this and/or any other of the examples cited, and the control or development of any particular patterns of movement in man. It is suggested, however, that study of other animals may permit isolation of some of the processes which interact to give the complex movement patterns found in man.

References

BERNSTEIN, N. (1967). *The Co-ordination and Regulation of Movements.* London Pergamon Press.

BULLOCK, T. H. (1961). The origins of patterned nervous discharge. *Behaviour*, **17**, 48–59.

DOTY, R. W. and BOSMA, J. F. (1956). An electromyographic analysis of reflex deglutition. *J. Neurophysiol.* **19**, 44–60.

DOTY, R. W., RICHMOND, W. H. and STOREY, A. T. (1967). Effect of medullary lesions on co-ordination of deglutition. *Exp. Neurol.* **17**, 91–106.

GIBSON, E. J. WALK, R. D. and TIGHE, T. J. (1959). Enhancement and deprivation of visual stimulation during rearing as factors in visual discrimination learning. *J. comp. physiol. Psychol.* **52**, 74–81.

GORSKA, T. and JANKOWSKA, E. (1961). The effect of deafferentation on instrumental (Type II) conditioned reflexes in dogs. *Acta Biol. exp., Vars.* **21**, 219–34.

GRAY, J. and LISSMANN, H. W. (1946*a*). Further observations on the effect of deafferentation on the locomotory activity of amphibian limbs. *J. exp. Biol.* **23**, 121–32.

GRAY, J. and LISSMANN, H. W. (1946*b*). The co-ordination of limb movements in the amphibia. *J. exp. Biol.* **23**, 133–42.

HASSENSTEIN, B. (1959). Optokinetische Wirksamkeit bewegter periodischer Muster. *Z. Naturf.* **14**, 659–74.

HASSENSTEIN, B. (1961). Wie sehen Insekten Bewegungen?. *Naturwissenschaften*, **48**, 207–14.

HASSENSTEIN, B. and REICHARDT, W. (1959). Wie sehen Insekten Bewegungen? *Umschau*, **10**, 302–5.

HEIN, A. and HELD, R. (1962). A neural model for labile sensori-motor co-ordinations. In *Biological Prototypes and Synthetic Systems.* I. New York: Plenum Press.

HEIN, A. and HELD, R. (1967). Dissociation of the visual placing response into elicited and guided components. *Science, N.Y.* **158**, 390–2.

HELD, R. (1961). Exposure-history as a factor in maintaining stability of perception and co-ordination. *J. nerv. ment. Dis.* **132**, 26–32.

HELD, R. and BAUER, J. (1967). Visually guided reaching in infant monkeys after restricted rearing. *Science, N.Y.* **155**, 718–20.

HELD, R. and HEIN, A. (1963). Movement-produced stimulation in the development of visually guided behaviour. *J. comp. physiol. Psychol.* **56**, 872–6.

HINDE, R. A. (1958). Alternative motor patterns in chaffinch song. *Anim. Behav.* **6**, 211–8.

HINDE, R. A. Ed. (1969). *Bird Vocalizations in Relation to Current Problems in Biology and Psychology.* Cambridge University Press.

HORRIDGE, G. A. (1966). Study of a system, as illustrated by the optokinetic response. *Symp. Soc. exp. Biol.* **20**, 179–98.

HORRIDGE, G. A. and BURROWS, M. (1968). Efferent copy and voluntary eye cup movement in the crab *Carcinus. J. exp. Biol.* **49**, 315–24.

HORRIDGE, G. A. and SANDEMAN, D. C. (1964). Nervous control of optokinetic responses in the crab *Carcinus. Proc. R. Soc. B.* **161**, 216–46.

HULSE, S. H. (1967). Licking behaviour of rats in relation to saccharin concentration and shifts in fixed-ratio reinforcement. *J. comp. physiol. Psychol.* **64**, 478–84.

IMMELMANN, K. (1969). Song development in the zebra finch and other Estrildid finches. In HINDE, R. A. (Ed.), *Bird Vocalizations in Relation to Current Problems in Biology and Psychology.* Cambridge University Press.

JAHN, T. (1960). Optische Gleichgewichtsregelung und zentrale Kompensation bei Amphibien, insbesondere bei der Erdkrote (*Bufo bufo* L.). *Z. vergl. Physiol.* **43**, 119–40.

KONISHI, M. (1963). The role of auditory feedback in the vocal behaviour of the domestic fowl. *Z. Tierpsychol.* **20**, 349–67.

KONISHI, M. and NOTTEBOHM, F. (1969). Experimental studies in the ontogeny of avian vocalizations. In HINDE, R. A. (Ed.), *Bird Vocalizations in Relation to Current Problems in Biology and Psychology*. Cambridge University Press.

LISSMANN, H. W. (1946*a*). The neurological basis of the locomotory rhythm in the spinal dogfish (*Scyllium canicula, Acanthias vulgaris*). I. Reflex behaviour. *J. exp. Biol.* **23**, 143–61.

LISSMANN, H. W. (1946*b*). The neurological basis of the locomotory rhythm in the spinal dogfish (*Scyllium canicula, Acanthias vulgaris*). II. The effect of deafferentation. *J. exp. Biol.* **23**, 162–76.

MACCORQUODALE, K. and MEEHL, P. E. (1954). Edward C. Tolman. In ESTES *et al.* (Ed.), *Modern Learning Theory*. New York: Appleton-Century-Crofts.

MACKINTOSH, N. J. (1964). Selective attention in animal discrimination learning. *Psychol. Bull.* **64**, 124–50.

MARLER, P. (1967). Comparative study of song development in sparrows. *Proc. 14th Int. Ornith. Cong. Oxford*, 231–44. Oxford: Blackwell.

MARLER, P. and TAMURA, M. (1964). Culturally transmitted patterns of vocal behaviour in sparrows. *Science, N.Y.* **146**, 1483–6.

MATTHEWS, P. B. C. (1964). Muscle spindles and their motor control. *Physiol. Rev*, **44**, 219–88.

MITTELSTAEDT, H. (1960). The analysis of behaviour in terms of control systems. *5th Cong. on Group Processes*. New York: Josiah Macy.

MITTELSTAEDT, H. (1962). Control systems of orientation in insects. *Ann. Rev. Entomol.* **7**, 177–98.

MITTELSTAEDT, H. (1964). Basic control patterns of orientational homeostasis. *Symp. Soc. exp. Biol.* **18**, 365–86.

NICOLAI, J. (1959). Familientradition in der Gesangsentwicklung des Gimpels (*Pyrrhula pyrrhula* L.) *J. Ornith.* **100**, 39–46.

NOTTEBOHM, F. (1967). The role of sensory feedback in the development of avian vocalizations. *Proc. 14th Int. Ornith. Cong. Oxford*, pp. 265–80. Oxford: Blackwell.

NOTTEBOHM, F. (1968). Auditory experience and song development in the chaffinch (*Fringilla coelebs*). *Ibis*, **110**, 549–68.

ROBERTS, B. L. (1969). Spontaneous rhythms in the motoneurons of spinal dogfish (*Scyliorhinus canicula*). *J. mar. biol. Ass.* In press.

SPERRY, R. W. (1950). Neural basis of the spontaneous optokinetic response produced by visual inversion. *J. comp. physiol. Psychol.* **43**, 482–9.

SPERRY, R. W. (1951). Mechanisms of neural maturation. In STEVENS, S. S. (Ed.), *Handbook of Experimental Psychology*. New York: John Wiley.

STEVENSON, J. G. (1967). Reinforcing effects of chaffinch song. *Anim. Behav.* **15**, 427–32.

STEVENSON, J. G. (1969). Song as a reinforcer. In HINDE, R. A. (Ed.), *Bird Vocalizations in Relations to Current Problems in Biology and Psychology*. Cambridge University Press.

SUTHERLAND, N. S. (1959). Stimulus analysing mechanisms. In *Proc. Sym. Mechanization of Thought Processes*, pp. 575–609. London: H.M.S.O.

SZÉKELY, G. (1968). Development of limb movements: embryological, physiological and model studies. In WOLSTENHOLME, G. E. W. and O'CONNOR, M. (Eds.), *Growth of the Nervous System*. London: M. Churchill.

TAUB, E., BACON, R. C. and BERMAN, A. J. (1965). Acquisition of a trace-conditioned avoidance response after deafferentation of the responding limb. *J. comp. physiol. Psychol.* **59**, 275–9.

TAUB, E., ELLMAN, S. J. and BERMAN, A. J. (1966). Deafferentation in monkeys: effect on conditioned grasp response. *Science, N.Y.* **151**, 593–4.

TEUBER, H. L. (1961). Sensory deprivation, sensory suppression and agnosia: notes for a neurologic theory. *J. nerv. ment. Dis.* **132,** 32–40.

THISTLETHWAITE, D. (1951). A critical review of latent learning and related experiments. *Psychol. Bull.* **48,** 97–129.

THORPE, W. H. (1961). *Bird-song.* Cambridge University Press.

THORPE, W. H. (1963). *Learning and Instinct in Animals* (1st Ed, 1956). London: Methuen.

WEIS-FOGH, T. (1964). Control of basic movements in flying insects. *Symp. Soc. exp. Biol.* **18,** 343–61.

VON HOLST, E. (1954). Relations between the central nervous system and the peripheral organs. *Br. J. Anim. Behav.* **2,** 89–94.

VON HOLST, E. and MITTELSTAEDT, H. (1950). Das Reafferenzprinzip. *Naturwissenschaften,* **37,** 464–76.

WELFORD, A. T. (1968). *Fundamentals of Skill.* London: Methuen.

WILSON, D. M. (1961). The central nervous control of flight in a locust. *J. exp. Biol.* **38,** 471–90.

WILSON, D. M. (1964). Relative refractoriness and patterned discharge of locust flight motor neurones. *J. exp. Biol.* **41,** 191–205.

WILSON, D. M. (1968). Inherent asymmetry and reflex modulation of the locust flight pattern. *J. exp. Biol.* **48,** 631–41.

WILSON, D. M. and WYMAN, R. J. (1965). Motor output patterns during random and rhythmic stimulation of locust thoracic ganglia. *Biophys. J.* **5,** 121–43.

WOLSTENHOLME, G. E. W. and O'CONNOR, M. (Eds.) (1968). *Growth of the Nervous System.* London: Churchill.

Q. Jl exp. Psychol. (1969) **21**, 127–133

ARE THRESHOLDS REDUCED BY ILLUSIONS? AN ATTEMPT AT REPLICATION

P. C. DODWELL, L. G. STANDING[†] AND HARRIET THIO
Department of Psychology, Queen's University, Kingston, Ontario

Three experiments are reported, which are attempts to replicate the finding of Ross and Gregory (1964) that difference thresholds for weights can be lowered by means of the size-weight illusion. Three different procedures were used, the first one (experiment I) being designed to show whether or not changes in a subject's judgement criterion could account for apparent changes in sensitivity. The second method (Experiment II) was a replication of Ross and Gregory's first procedure, in which the standard weight was judged before the comparison. In Experiments I and II a larger illusion was induced than in the original studies, but in Experiment III both the weights and container sizes were practically identical to those used by Ross and Gregory. The procedure was also the same as their most successful procedure (number 3) in which standard and comparison were judged simultaneously. The findings were uniformly negative: there was no evidence of criterion shift when the size-weight illusion was induced nor did we find the lowering of threshold previously reported.

Introduction

Ross and Gregory (1964) report that sensitivity for weight differences can be increased simply by altering the phenomenal weights of objects by means of the size-weight illusion. They compared differential thresholds by the frequency method for a standard weight of 200 g. (a) when judgements were made using 35 mm. film tins ($1\frac{1}{4}$ in. × $1\frac{3}{4}$ in.), and (b) when judgements were made using Nescafé tins (3 in. × 3 in.). With their most successful procedure, when standard and comparison weights were judged simultaneously, they found a reliable decrease in the difference threshold under (b) the phenomenally lighter set.

This finding, as the authors themselves point out, is quite surprising. For a hundred years or so thresholds have been regarded as defining the upper limits of sensitivity, limits which could not be exceeded by experimental manipulation. Although the advent of the theory of signal detection (Swets, Tanner and Birdsall, 1961) cast doubt on the concept of threshold, probably most psychologists subscribe to the view that the upper limit of sensitivity is set by inherent factors in the organism, such as neural "noise." It therefore seems paradoxical that sensitivity might be increased by an experimental factor, particularly if this involves stimulation of a different modality as in the size-weight illusion. Moreover there is a logical puzzle involved in appealing to Weber's law to account for greater differential sensitivity to apparently lighter objects, as Ross and Gregory do. If the mechanism which increases sensitivity really is reduction of the phenomenal magnitude of the standard weight, the same factor will also presumably affect the phenomenal magnitude of the comparison weight increment which is to be detected, thus cancelling out rather precisely the postulated increase in sensitivity if Weber's law holds. Finally, Cornsweet and Teller (1965) have shown that increment thresholds for luminance are unaffected by changes in the apparent brightness of the background field, but depend only on its luminance. Here we have a situation where the physical parameters of the stimulus display clearly determine the detectability of stimulus change, although within a single modality and in a very different situation.

[†] Present address: University of Rochester, Rochester, N.Y., U.S.A.

The effect of apparent weight on sensitivity clearly merits further study, particularly since Ross and Gregory's results were not clear cut. They report one reliable result, one which is somewhat marginal, and one which is non-significant, on the basis of three different experiments using the same weights but slightly different procedures. Two "unsuccessful" experiments are also mentioned. An alternative explanation for their positive findings suggests itself; since they used conventional psychophysical methods, any change in sensitivity was necessarily confounded with possible changes in an observer's criterion for making a particular judgement. In signal detection terminology (Swets *et al.*, 1961) a change in β without any increase in d′ (the measure of sensitivity) could lead to increased report of weight differences. But whether a change in detection of weight differences represents a genuine change in sensitivity would depend on whether, and to what extent, the frequency of false detections (false alarms) also changes.

Three experiments were done. The aim of the first was to test the hypothesis that an apparent decrease in the difference threshold for phenomenally lighter weights, if found, could be attributed to changes in a subject's response criterion rather than change in sensitivity to weight differences. The procedure of Ross and Gregory was therefore modified to allow the estimation of both hit (correct identification) and false alarm rates. In the second experiment their first procedure (which yielded a just significant result) was closely followed, and in the third experiment their third and most successful procedure was fully replicated.

Experiment I

Method

Subjects. Sixty-four experimentally naive subjects from an introductory psychology class at Queen's were used. None of them was aware of the purpose of the experiment.

Apparatus. Two sets of weights were employed. Each consisted of two 200 g. weights, together with 6 heavier and 6 lighter weights (differing by plus or minus 5, 10, 15, 20, 30 and 40 g. from the standard). One was a set of small metal cylinders (6 cm. high, 3 cm. in diameter), the other a set of large cylinders (21 cm. high, 10 cm. in diameter). All stimuli were painted black, and within a set were identical in appearance. They were lifted by means of rings, attached to the cylinders with thread, to avoid problems of difference in finger span between the two sets, just as in Ross and Gregory's method.

Procedure. The subjects were randomly assigned, eight to each of the eight treatment combinations of a 2 × 2 × 2 factorial design (large vs. small cylinders, positive vs. negative weight increments, and ascending vs. descending order of testing with the successive magnitudes of weight increment). The experimental task involved the forced-response discrimination between presence or absence of a weight difference when a test stimulus and the 200 g. standard stimulus were lifted simultaneously by the subject. On 50 per cent of the trials, chosen randomly, the test stimulus also weighed 200 g. Twenty trials (including the blank trials) were given for each of the 6 weight increments (or decrements) before the next largest increment or decrement was employed. Thus, each subject received 120 trials.

Before being tested, the subject was told that there would be a weight difference between the two stimuli on only 50 per cent of the trials; he was also told whether the test stimulus would be heavier or lighter than the standard when there was a difference. The subject was instructed always to lift the two stimuli simultaneously (by far the most reliable of Ross and Gregory's three procedures for lifting the weights), one with each thumb and forefinger, and to lift the standard stimulus with the right hand. The subject was required to respond on all trials, and responded "same" or "heavier"/"lighter." The subject was also required to give a confidence rating on each trial, from 1 (least confident of his decision) to 5 (most confident). Feedback was provided after each trial by the experimenter saying "correct" or "incorrect."

A demonstration of each weight increment, and of a blank trial, were given before the block of 20 trials on which that increment was employed.

Results

The results definitely do not show any superior discrimination of weight by the subjects tested with the large (and thus phenomenally lightened) cylinders, since the total number of correct discriminations by all subjects under this condition was slightly *lower* than for the subjects receiving the small (and phenomenally heavier) stimuli: 2659 as compared with 2698 correct responses. An analysis of variance was performed on the results to determine the effects on the probability of correct response of increment magnitude, positive versus negative increments, order of testing, and small versus large cylinders. The only significant factor was the magnitude of the weight increment employed ($P<0{\cdot}01$). A second analysis showed that variability of performance was not affected significantly by any of the above factors. The probability of correct response is shown in Figure 1 for each weight increment, for the large and the small stimuli.

Other aspects of the subjects' performance were also found to be unaffected by the change from small to large cylinders. The numbers of hits and false alarms were 1284 and 562, respectively, for the large stimuli, compared with 1358 and 562 for the small stimuli. Confidence ratings on correct and incorrect trials totalled 9505 and 3480, respectively, for the small stimuli; for the large stimuli the corresponding totals were 8590 and 3347. In only one case did individual comparisons of the difference between the large and small stimulus conditions even approach statistical significance at the 0·05 level.

Experiment II

Failure to demonstrate a change in sensitivity in experiment I might have been a function of the change to TSD procedures; also the dimensions of the larger set of weights were quite a bit bigger than of the large set used by Ross and Gregory, thus giving a greater illusion (about 30 per cent, as against their 20 per cent). To decide whether the former possibility might hold, in Experiment II Ross and Gregory's first procedure was followed, but using the same sets of weights as

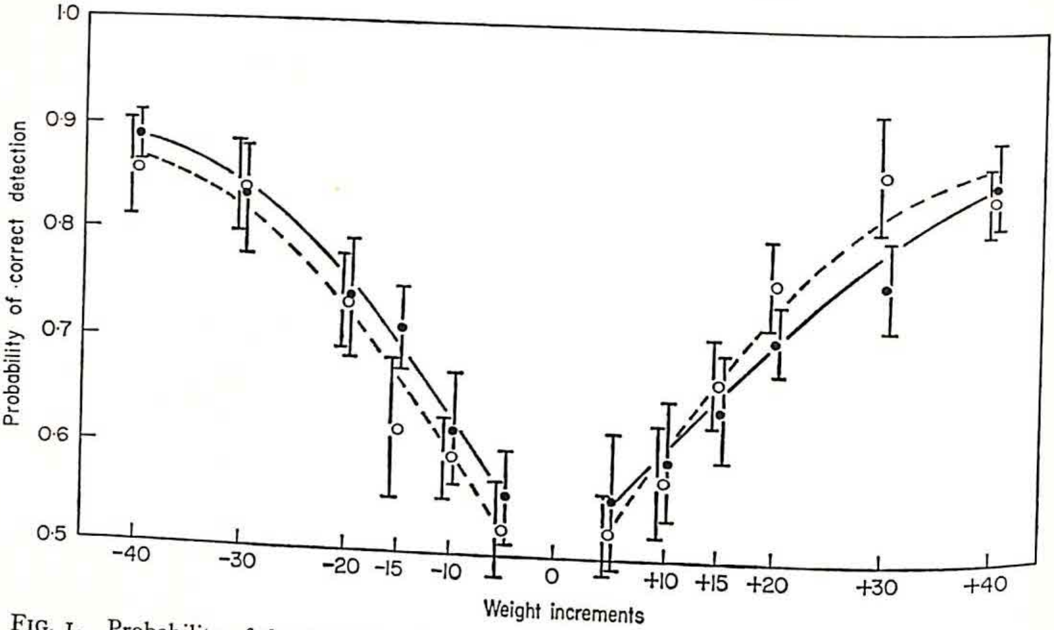


FIG. 1. Probability of the detection of a weight difference as a function of the actual difference (Experiment 1). Each data point is the mean of 160 readings. ●, Small stimuli; ○, large stimuli.

in our Experiment I. By this means we hoped to find out whether replication of their first procedure would yield more reliable results, since one would expect the greater illusion to generate a somewhat bigger change in threshold than in the original study.

Methods

Subjects. Sixteen nurses at the Kingston Psychiatric Hospital served as volunteer subjects. Their ages were all between 19 and 21 years, and all were naive with respect to the purpose of the experiment.

Procedure. The same sets of weights were employed as in Experiment I. The subject was seated at a table with one of the standards in front of her. The test weights were placed in turn beside the standard in a predetermined random order. All subjects were tested on both sets of weights, half with the larger set first, and half with the small set first. The following instructions were read to each subject at the start of the session: "This experiment involves judging weights. Each time you will be presented with 2 weights. You are to lift these weights with your preferred hand by holding the loops between your thumb and forefinger, lifting smartly just once, like this (demonstration). This weight will be in front of you all of the time. You are to lift this standard weight first on each trial, and the other weight second. You are to indicate whether the second weight is lighter or heavier than the first weight. If you are not sure whether a weight is heavier or lighter than the standard, please make the best guess you can." Each comparison was made ten times, yielding 120 judgements for each of the two sets.

Results

Wilcoxon's matched-pairs signed-ranks test (Siegel, 1956) was used to determine the reliability of the difference between the points of subjective equality (PSEs) and standard deviations (S.D.s) as calculated by the method of averaged *z* scores (Woodworth and Schlosberg, 1955, p. 205). The data for individual subjects are shown in Table I, together with means and medians. The difference in PSE

TABLE I

Experiment II: standard first, comparison second

Subject	PSE Large	PSE Small	S.D. Large	S.D. Small
1	196·6	194·3	11·6	12·2
2	194·3	193·8	17·0	13·9
3	191·8	200·9	13·6	12·6
4	198·5	189·8	14·5	10·1
5	202·0	192·2	12·4	12·9
6	181·4	185·3	16·8	11·6
7	203·6	203·9	12·3	11·9
8	193·6	191·7	11·2	11·2
9	194·4	184·8	10·6	11·7
10	192·5	197·6	13·8	16·5
11	201·3	193·1	13·1	14·5
12	195·1	193·3	12·6	13·7
13	192·9	188·9	23·2	17·5
14	183·9	187·0	18·6	15·9
15	203·0	188·2	12·1	12·0
16	190·8	188·1	11·1	13·1
Median	194·35	191·95	12·85	12·75
Mean	194·73	192·06	14·03	13·21

The differences between PSEs and between S.D.s are not significant.

for the two sets was in the same direction as in Ross and Gregory's experiment, but was not reliable. The difference in S.D.s was also not reliable, but the actual difference was in the opposite direction to that reported previously, i.e. the S.D. for the larger set tended to be greater than for the small set, thus indicating a higher threshold, if anything. However, the evidence is not sufficient to reject the hypothesis of no difference in sensitivity for the two sets.

Experiment III

In order to leave no doubt about the replicability of threshold changes, Ross and Gregory's Experiment III was repeated as closely as possible.

Method

Procedure. The procedure was the same as in our Experiment II, with appropriate modifications in instructions, so that standard and test weights were lifted simultaneously. Sixteen undergraduates and graduate students served as volunteer subjects, and had no knowledge of the purpose or expected outcome of the experiment. Two new sets of weights were made, exactly the same sizes as Ross and Gregory's. Instead of a thread and ring, each cylinder had a stiff wire semicircular handle rigidly attached over its top, care being taken to make sure no differential visual cues were available.

Results

The results (calculated as in Experiment II) are shown in Table II. Again there is no reliable difference between PSEs or S.D.s, as assessed by the Wilcoxon test, although the mean difference in S.D. is in the same direction as in our second experiment, that is, against the trend reported by Ross and Gregory.

TABLE II

Experiment III: simultaneous

Subject	PSE		S.D.	
	Large	Small	Large	Small
1	200·2	202·6	11·2	9·2
2	192·5	199·6	11·4	10·8
3	210·9	203·0	9·8	10·0
4	190·4	191·7	17·3	13·9
5	193·6	192·4	11·8	15·4
6	200·7	192·8	9·5	10·8
7	188·4	191·9	15·5	17·8
8	194·9	196·4	20·0	32·2
9	201·2	201·3	9·1	11·2
10	206·0	202·0	14·4	8·6
11	191·4	191·2	8·5	14·4
12	203·3	198·2	13·0	14·0
13	206·9	200·1	10·4	9·8
14	209·2	207·4	11·8	12·7
15	200·4	203·5	13·3	17·6
16	184·2	187·0	15·6	13·6
17	198·1	199·0	14·0	9·9
18	196·0	196·6	11·7	10·1
19	196·0	200·0	8·8	11·5
20	206·3	212·3	13·2	13·3
Median	199·1	199·3	11·8	12·1
Mean	198·1	198·5	12·5	13·3

The differences between PSEs and between S.D.S are not significant.

Discussion

We conclude from these three experiments that there is no effect on sensitivity to differences in lifted weights which can be attributed to changes in apparent weight brought about by means of the size-weight illusion. Experiment III in particular failed to demonstrate a change in threshold for weight differences under circumstances which were nearly identical to those of Ross and Gregory's Experiment III, where the effect was very reliable ($P < 0·005$). The only differences between these two experiments, so far as we know, were first that we used a slightly larger range of weights (a range of 80 g. in each set, as against their range of 50 g.) which if anything should have increased the precision of our estimates, since even at the extremes of the range subjects were not making 100 per cent correct judgements (see Fig. 1). Second, although our objects were almost exactly the same size as theirs, ours had rigid handles rather than the thread and ring arrangement. It is difficult to believe that this could be a crucial difference. Third, we used different subjects, although they were probably fairly comparable in terms of age and other characteristics. However, it is possible that in one important respect some of their subjects did differ from ours, namely in their knowledge of the purpose and expected outcome of the experiment. Unless care was taken to keep them uninformed about this, it could be that the judgements of their undergraduates and research students were influenced by such knowledge. It is known that knowledge and a set to respond in a particular way can influence judgements in a psychophysical type of situation (e.g. Gaze and Dodwell, 1965). Apart from this suggestion, we can think of no way of reconciling the two sets of findings except the obvious one, that Ross and Gregory's two successful experiments were genuine cases of type I error. It can scarcely be argued that we failed to find a true difference because our data are more variable than theirs: for their Experiment I, the range of individual S.D.s is 5·9 and 14·2 for large and small sets, respectively, and in our comparable Experiment II the corresponding ranges were 12·6 and 7·4. In their Experiment II the ranges were 11·6 and 12·5, in our Experiment III they were 11·2 and 23·6, or excluding one anomalous subject (no. 8, Table II) for the small set, also 11·2.

Originally we had expected there might be criterion changes in judging the larger sets of weights, but our Experiment I failed to show any. It is conceivable, if Ross and Gregory's subjects knew the hypothesis which was being tested, that they did shift criterion, and it might be worth finding out whether this occurs with knowledgeable subjects. We emphasize again that with the methods they used, there is no means of knowing whether the apparent change in threshold they obtained was due to this factor.

We conclude that there is no evidence that Weber's law applies to the apparent rather that the physical values of the stimuli, at least for weights, and in view of Cornsweet and Teller's result (*op. cit.*, 1964) we think it highly probable that the statement can be generalized to other modalities.

This work was supported by grant 9425–12 of the Defence Research Board of Canada to the first author which is gratefully acknowledged. We also thank Gail Eidinger and Ruth Belch for assistance in data collection and analysis.

References

Cornsweet, T. N. and Teller, D. Y. (1965). Relation of increment thresholds to brightness and luminance. *J. opt. Soc. Am.* **55,** (10), 1303–8.

Gaze, L. and Dodwell, P. C. (1965). The role of induced set in figural after-effects. *Psychon. Sci.* **2,** 275–6.

Ross, H. E. and Gregory, R. L. (1964). Is the Weber fraction a function of physical or perceived input? *Q. Jl exp. Psychol.* **16,** 116–21.

Siegel, S. (1956). *Non-parametric Statistics.* New York: McGraw-Hill.

Swets, J., Tanner, W. P. and Birdsall, T. G. (1961). Decision processes in perception *Psychol. Rev.* **68,** 301–40.

Woodworth, R. S. and Schlosberg, H. (1955). *Experimental Psychology.* New York: Holet.

Received 24 *October* 1968

Q. Jl exp. Psychol. (1969) **21**, 134–136

PRIOR STIMULATION AND THE PERCEPTION OF CAUSALITY

H. V. BROWN AND T. R. MILES

Psychology Laboratory, University College of North Wales, Bangor

An attempt was made to discover whether, in conditions similar to those of Michotte, the number of causal responses given by subjects was influenced by stimulation during the period just before the main experiment. Time-intervals between the arrival of object A and the departure of object B were systematically varied within the range 60 to 390 msec. Those who had earlier been exposed to shorter time-intervals (60 to 210 msec.) gave fewer causal responses than those who had been exposed to longer ones (150 to 300 msec. and 240 to 390 msec.). A possible explanation is offered in terms of the modification of existing "schemata."

Introduction

Michotte (1963, pp. 91 *et seq.*) claims that causal responses disappear when the time-interval between the arrival of object A at the point of impact and the departure of object B exceeds 200 msec. It now seems established in all experiments of this kind that individuals differ considerably in what they report (Beasley, 1967) and that previous experience can affect the result in a variety of ways (Grüber, Fink and Damm, 1957; Powesland, 1959; Crabbé, 1967). The present experiment was designed to study the effect of exposure to different time-intervals in the period immediately before the main experiment.

Method

Apparatus, subjects and procedure

The apparatus was basically similar to that described by Michotte (1963, pp. 27 *seq.*). It consisted of a rotating disc mounted on a fast-stepping motor, the speed of which was controlled by electronic pulses. Immediately in front of the disc was a screen, with a slit in it 15 cm. in length and 0·5 cm. wide. Rotating with the disc were two pieces of cardboard on both of which were drawn two strips in accordance with the mathematical formulae described by Michotte (1963, p. 28). One strip was blue, the other black (Michotte in fact used red and black, but we decided against red because of possible complications if any subjects were colour blind). The upper piece of cardboard could be adjusted in relation to the lower to give the required time-interval between the arrival of A and the departure of B. The observation distance was 1 to 1·5 m. Viewing conditions were those of normal daylight, except that the slit was illuminated by two small lamps behind the screen. Fifteen subjects were used, 10 of them being first year Psychology students, the other 5 belonging to other departments but having had a small amount of experience in psychological experimentation.

Twelve different time-intervals (between the arrival of A and the departure of B) were used, ranging from 60 to 390 msec. at intervals of 30 msec. In both the preliminary and main experiments all subjects were restricted to three categories of response, viz. "A" responses (Michotte's "direct launching"), "B" responses ("delayed launching"), and "C" responses ("two separate movements"). These three categories were explained verbally to the subjects, who were then handed a card reminding them of the essential features. There were three groups of subjects. In the preliminary part of the experiment, group S ("short") were exposed to movements separated by time-intervals of 60, 90, 120, 150, 180 and 210 msec., group M ("medium") to time-intervals of 150, 180, 210, 240, 270 and 300 msec., and group L ("long") to time-intervals of 240, 270, 300, 330, 360 and 390 msec. Each group contained 5 subjects, and the 6 time-intervals were presented in random order except for the stipulation that each should occur twice and that the same one should not occur succesively. When this part of the experiment had been completed there was a break of 2 min. during which neutral topics were discussed. In the main experiment all 12 time-intervals (60 to 390 msec.) were presented, again in random order except for the stipulation that each time-interval should occur twice and that the same one should not occur successively. Since there were 15 subjects this gave a "pool" of 15 × 24 (360) readings. Results are given in Table I.

TABLE I

Number of responses in categories "A," "B" and "C" for the three groups of subjects

	Response-category		
	A	B	C
Group S	39	51	30
Group M	50	63	7
Group L	43	68	9

Discussion

Table I shows that those subjects who had been exposed to shorter time-intervals at the preliminary stage gave fewer causal responses in the main experiment.†

This result supports the claim made both by Grüber *et al.* (1957) and by Powesland (1959) that in experiments on the perception of causality prior stimulation can influence the result. More specifically it establishes that, if the conditions are standard in other respects, the occurrence or non-occurrence of causal responses is not simply a function of the absolute value of the time-interval separating the movements, as Michotte's figures (1963, pp. 91 *et seq.*) seem to suggest.

Two explanations of our result seem possible. The first is that the subjects in group S underwent a change in adaptation-level, of the kind described by Helson (1964). In other words, as a result of the prior stimulation the longer time-intervals in the main part of the experiment actually seemed longer to these subjects than they otherwise would have done. Alternatively it is possible that what was involved was a "semantic adjustment" of the kind described by

† χ^2 for the table as a whole comes to 25·10 ($P<0{\cdot}001$). If A and B responses are pooled, and the results for groups M and L are pooled the resultant 4-cell table (causal vs. non-causal responses, subjects exposed to short time-intervals vs. others) gives a χ^2 of 24·12 ($P<0{\cdot}001$). As a check one may pool the B and C responses, contrasting them with the A responses; if the results for groups M and L are pooled as before, χ^2 comes to 1·35 (n.s.). The significant factor is thus shown to lie in the C responses. It is perhaps worth adding that one subject in group S gave no C responses at all, and that the large number of C responses for group S occurred in spite of this. The problem of individual differences, however, is not one which we have considered in this paper.

Stevens (1958, p. 634)† In other words these subjects altered their standards of what was to merit the term "causal." The first explanation does not call for any major modification of Michotte's views; he could simply say that the presence or absence of the causal impression was a function of "phenomenal" (or "felt") time rather than of "physical" time. If, however, people's standards of what time-interval justifies a causal response can be modified, this seems hard to square with Michotte's claim (pp. 21–22) that a particular phenomenal datum ("*the* causal impression") is dependent on specifiable stimulus conditions.

Although our results do not distinguish between a "change-in-adaptation-level" hypothesis and a "semantic adjustment" hypothesis (since it was not possible to combine our experimental procedure with requests for assessment of the length of temporal intervals), it is perhaps of interest to suggest a new theoretical formulation which both accommodates Michotte's main findings and is also compatible with the "semantic adjustment" hypothesis. A further advantage of this formulation is that all reference to the controversial notion of a "causal impression," in Michotte's sense, is avoided.

The view which we are advocating involves the notion of "schemata," in the sense given to this word, e.g. by Oldfield (1941, pp. 42–5). Past experience, we suggest, gives people schemata for likely time-intervals between cause and effect in situations where they know that physical work is being done. Michotte's great achievement was to provide a technique for presenting the requisite visual stimuli under "illusory" conditions, i.e. in the absence of physical work. If in these conditions an experimenter systematically varies the time-interval between the two movements and then asks subjects if they had a "causal impression," he is in effect saying, *Does it still look as if the time-intervals are right*? Our result would in that case entitle us to claim that stimulation of a certain kind before the main experiment can sometimes modify previously formed schemata.

References

Beasley, N. D. (1967). Unpublished Ph.D. thesis. University of Nottingham.

Crabbé, G. (1967). *Les Conditions d'une Perception de la Causalité.* Monographes Francaises du Psychologie, XII, Paris.

Grüber, H. E., Fink, C. D. and Damm, V. (1957). Effects of experience on perception of causality. *J. exp. Psychol.* **53,** 89–93.

Helson, H. (1964). *Adaptation-Level Theory.* New York: Harper and Row.

Michotte, A. (1963). *The Perception of Causality.* London: Methuen.

Oldfield, R. C. (1941). *The Psychology of the Interview.* London: Methuen.

Powesland, P. F. (1959). The effect of practice upon the perception of causality. *Can. J. Psychol.* **13,** 155–68.

Stevens, S. S. (1958). Adaptation-level vs. the relativity of judgement. *Am. J. Psychol.* **71,** 633–45.

Received 24 October 1968

---

† Stevens cites the example of someone who "call(s) a 400-g. weight 'very heavy' when he lifts it along with lesser weights, and . . . call(s) it 'medium' when he lifts it in a series that includes a weight of 900 g.," even though there is no change in the sensory character of what is perceived.

Q. Jl exp. Psychol. (1969) **21**, 137–147

GROUPING AND SHORT-TERM MEMORY: DIFFERENT MEANS AND PATTERNS OF GROUPING

JOANNA RYAN[†]
Psychological Laboratory, Cambridge

Two experiments, concerned with the improving effects of grouping on auditory short-term memory, are described. In the first, temporal grouping was found to improve recall considerably, but non-temporal grouping had a much smaller effect. Temporal grouping reduced the order errors more than other errors; it also changed the pattern of the order errors. Further, it altered the shape of the serial position curve of all errors. In the second experiment, irregular patterns of temporal grouping were found to be inferior to a regular pattern. The results are discussed in terms of the time available for processing previous items during the presentation of a sequence, and the form that this processing may take.

Introduction

It is a commonplace observation that recall is improved when the material to be remembered is in some way organized or grouped. In short-term memory tasks several means of grouping—temporal, rhythmical, spatial, imposed either externally or by the subject himself—have been studied; the most usual finding is an improvement in recall with grouping (e.g. Adams, 1915; Conrad and Hille, 1957; Michon, 1964; Wickelgren, 1964). The extent of this improvement depends on the size of the groups. Thus, grouping into groups of two items each is generally not helpful, whereas both grouping into threes and grouping into fours usually are helpful; grouping into fives and into sixes have only sometimes been found helpful, usually with longer sequences of more than ten items (e.g. Pollack, Johnson and Knaff, 1959; Wickelgren, 1964; Mayzner, Tresselt, Adler, Cohen and Schoenberg, 1966).

Apart from such investigations of different group sizes, few other aspects of grouping have been studied, and very little theoretical interest has been shown in it. Grouping does, however, raise some important questions about short-term memory mechanisms. In general, an explanation is required of how it can cause an improvement in capacity. Miller's (1956) ideas of recoding and "chunking" have an obvious, if rather vague, relevance to this: the items of one group might be recoded to form one chunk, instead of several separate items, with the consequent advantages suggested by Miller. In order, however, to specify the precise system of recoding that grouping induces, it is necessary to know in detail both how the various means and forms of grouping vary in their effectiveness, and also whether grouping causes any other changes in recall, apart from the overall improvement.

[†] Present address: Unit for Research on Medical Applications of Psychology, Station Road, Cambridge.

The present concern is with temporal grouping: with sequences of nine items presented auditorily and grouped by slight pauses within each sequence into three groups of three items. The question investigated is how far the temporal aspect of this grouping, that is, having the groups demarcated by pauses, is essential to its effectiveness. It is conceivable that non-temporal grouping, with the same mode of presentation, pattern of grouping, and sequence duration, but with different means of demarcation, would be as effective as temporal grouping. However, with temporal grouping the inter-item rate of presentation varies within each sequence, whereas with non-temporal grouping (or with no grouping at all) this rate is constant throughout each sequence. It is possible that any beneficial effects of temporal grouping are due to this variation in rate, rather than simply to the division of the sequence into subsections. This is particularly likely if what the subject does during the presentation of a sequence has an important influence on his subsequent recall. For example, besides trying to attend to the incoming items, the subject may also be trying to preserve those already presented. If so, then the precise distribution in time of the incoming items will determine the amount and the nature of the processing of previous items that can be carried out, and also how the subject alternates between the two tasks of attending to new items and preserving old ones. In a temporally grouped sequence the pauses may allow time for more uninterrupted processing (where this can mean rehearsal, coding, organizing), with less frequent alternation, than is available in an ungrouped or non-temporally grouped sequence. If this is the case, then temporal grouping should be superior to non-temporal grouping, and the latter should not differ from no grouping.

It might, on the other hand, be argued that it is the division alone which makes a sequence easier to handle, and that any means of demarcating the groups should be equally effective as long as such division is achieved. Such a view gains some support, at least by analogy, from the notion of "perceptual anchors," as used by Garner (1962), in a discussion of absolute judgement experiments. Garner argues that the provision of anchors or reference points within a series of stimuli should improve discrimination by, in effect, presenting the subject with several separate ranges of stimuli instead of one, within each of which normal discrimination can be achieved. Grouping may well be helpful because it operates in a manner analogous to such anchors, dividing the continuum of sequence positions, and thus allowing better discrimination of the position of each item. If the pauses in a temporally grouped sequence do act only as divisive anchors in this sense, then any other means of division should be as helpful. Further, it would on this analogy be expected that grouping, where it was effective, would particularly benefit the recall of the order of items, as opposed to recall of which items were included in the sequence.

Apart from the comparison of temporal with non-temporal grouping the following experiment was also done with the aim of providing more detail than is so far available about the effects of grouping on recall errors.

Experiment I

Procedure

Grouping methods

Grouping into three groups of three, with nine-digit sequences, was chosen, since it had been found previously (Ryan, 1967) that this gave both a convenient number of errors in the ungrouped condition and a sizeable improvement with grouping. The chief method used to provide non-temporal grouping was to interleave brief pips between the appropriate items of an otherwise ungrouped sequence. This particular method was chosen since it demarcates the groups by the interpolation of an event between the items without altering the items themselves, and in this respect resembles the way pauses demarcate groups. Possible non-temporal methods such as change of voice or inflection were not used because they would demarcate the groups by altering the nature of the items. Two additional non-temporal conditions were used, in order to control for the possibility in the main non-temporal condition that the pips would only suggest a strategy of grouping as distinct from facilitating it. In one condition no external indication of grouping was given, but the subjects were instructed to group in threes. In the other, both pips and instructions were used. This latter also formed a control for the possibility that the pips might have interfering effects.

Design

There were five conditions: ungrouped, UG; temporally grouped, TG; pips only, P; instructions only, I; and pips plus instructions, P + I. There were 50 subjects drawn from Cambridge University; ten different ones were tested in each condition. Each condition consisted of the same forty sequences, in the same order, except for the ungrouped condition. This consisted of 56 sequences, of which the first 40 were the same as those in the other conditions. Subjects were assigned randomly to all the grouping conditions; the assignment to the ungrouped condition was done separately, a few months earlier.

Sequences

The sequences were compiled by taking numbers from a random number table with the following constraints: there were no repetitions of any one digit; no digit could occur next to a numerically adjacent one; obvious patterns (e.g. 2, 4, 6, 8) were not allowed; over all the sequences the digits occurred with approximately equal frequency in each serial position.

Rates of presentation

For all five conditions the duration of each sequence was the same, namely 5 sec. This resulted in a uniform presentation rate of 1 digit every 0·55 sec. for the ungrouped and non-temporal conditions. In the temporally grouped, condition within each group, the items were presented at the faster rate of 1 digit every 0·45 sec.; the pauses between groups (from the beginning of the last item of one group to the beginning of the first item of the next) lasted 0·9 sec. The interval between sequences was 10 sec. The pips were recorded separately on to the second channel of the tape recorder, so that they occurred after every third digit of the appropriate sequences. The duration of each pip was about 0·2 sec. and all the sequences were checked to ensure there was no masking of the digits when the two channels were played back together.

Instructions

Subjects were tested individually, with spoken recall. In all conditions subjects were informed about the length and constitution of the sequences. They were told to repeat each sequence as quickly as possible, and always to guess if in doubt. In the ungrouped condition they were told that the digits would occur at a uniform rate; in the three non-temporal conditions they were also told this. In addition, in the pips only condition (P) they were told that a brief pip occurred after every third digit; in the instructions only (I) condition they were instructed to "try to group each sequence into groups of three, to try to break up each sequence into threes." In the combined condition (P + I) they were given the same instructions about grouping, and they were further told about the pips; it was then pointed out that these might be helpful in carrying out the instructions. In the temporally grouped condition they were told that each sequence was divided into three groups of three,indicated by slight pauses. One example of the appropriate kind of sequence was played to each subject, and five practice sequences were then given.

Scoring

The sequences were scored by counting the individual errors in each sequence, for each subject. Errors were categorized according to the loss they represented from each serial position, as follows; (i) transposition—any item not in its correct serial position, but occurring elsewhere in the sequence, and (ii) item error—any item neither in its correct position, nor included elsewhere in the sequence. These two categories cover all the possible kinds of error *losses*. This method of scoring was used in preference to the more usual one, in which errors are categorized according to what replaces a mistaken item, mainly for its simplicity (see, for example, Conrad, 1959, and Murray, 1965, for the great variety, ambiguity and incompleteness of categories of replacing errors that can be used). A separate account was, however, taken of the number of omissions, that is, instances where there is no replacing item.

Results

(i) *Overall performance.* The mean overall error scores (combining the different categories of errors) per subject, per sequence, for each condition, are shown in Table I. It can be seen that condition TG caused a much larger reduction in errors, compared to condition UG, than did any of the non-temporal conditions. The difference in mean scores across conditions was found to be significant using a Kruskal-Wallis (Siegel, 1956) one-way analysis of variance ($H = 11$, $d.f. = 4$, $0{\cdot}05 > P > 0{\cdot}02$). The differences between individual conditions were assessed using the Mann-Whitney U-test (two-tailed). Only the following comparisons gave significant results (values of $P \leqslant 0{\cdot}05$ being required throughout): conditions UG and TG ($U = 6$, $P < 0{\cdot}002$); conditions TG and P($U = 13$, $P < 0{\cdot}02$).

TABLE I

Mean error score, per sequence and per subject

Ungrouped† UG	Pips only P	Instructions only I	Pips + instructions P + I	Temporal grouping TG
2·11	1·76	1·58	1·41	0·55

† The mean for condition UG was calculated from the first 40 of the 56 sequences used.

(ii) *Serial position curves.* Figure 1 shows the total number of all errors at each serial position, for each grouping condition, and the total for the first 40 sequences of the ungrouped condition. The shape of the curve for condition UG resembles that commonly found in many learning tasks, that is, negatively skewed and

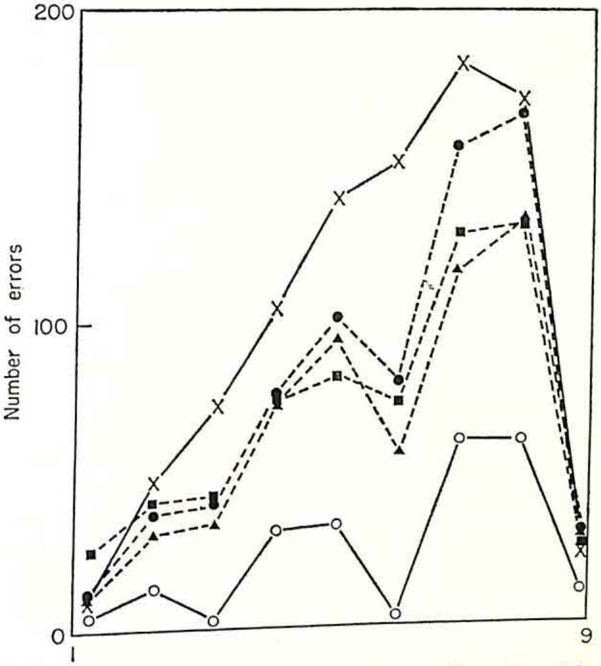


FIG. 1. Serial position curves, -X—X-, ungrouped;-●- - -●-, pips; -■- - -■-, instructions, --▲- -▲-, pips and instructions; —○—○—, temporally grouped.

bowed. Condition TG, apart from the overall lowering of the curve, also changes its shape, with dips at the third and sixth positions. Similar but smaller changes are found in the non-temporal conditions. Mann-Whitney U-tests were performed for each serial position separately. Conditons UG and TG showed significant differences for all serial positions; conditions UG and each of the non-temporal conditions showed significant or nearly significant ($0{\cdot}1 > P > 0{\cdot}05$) differences at position 6 only.

(iii) *Error analysis.* The mean proportion that each kind of error formed out of the total number of errors was found, for all subjects. Omissions were between 1 and 3 per cent of the total for all conditions, indicating that the instructions not to omit were followed successfully. In condition UG transpositions formed 68·6 per cent of the total; in condition P, 61·8 per cent; in condition I, 63·3 per cent; in condition P + I, 78·5 per cent, and in condition TG, 51·5 per cent. The transpositions were analysed further according to the distance of transposition—that is, adjacent, two-apart, three-apart, etc. The frequency of each distance of transposition for each of the conditions is shown in Table II. In condition UG the frequencies of these different transpositions bear an inverse relation to their distance, i.e. the nearest are the most frequent. This is the general form of the relationship that would be expected by chance since there are more possible transpositions involving close items than distant ones; it is not however the exact relationship that would be expected, since relatively many more close transpositions, and many fewer distant ones are found. In condition TG this inverse relationship is altered: as well as the overall reduction in transpositions it can be seen that two-apart ones are now less frequent than those three-apart,

TABLE II

Frequency of transpositions according to distance of displacement

Condition	Adjacent	Distance of displacement 2-apart	3-apart	4-apart	5-apart	6-apart	Total
UG	371	140	67	30	—	—	608
P	252	80	89	10	3	3	437
I	201	76	87	18	5	3	390
P + I	235	92	89	14	8	4	442
TG	63	10	36	1	3	—	113

and that the latter have increased their frequency relative to adjacent errors. The three non-temporal conditions appear to have an intermediate effect on the distance of transpositions.

Discussion

Considering temporal grouping alone first, this experiment has shown how in many respects it changes the pattern of errors made in recall, as well as causing a significant reduction in these. The qualititive effects—the change in the shape of the serial position curve, the decrease in the percentage of transposition errors and the change in the relative frequency of these—can all be used as indications of the kind of recoding that grouping induces. Only some general points will be made about this here: first, the error patterns reflect to some extent the structure of the imposed grouping; and second, temporal grouping gains most of its advantage from an improvement in the recall of the order of the items within a sequence, as opposed to the content alone of the sequence. The validity of this latter claim depends in turn upon the validity of distinguishing between two forms of information—information about the occurrence of items, and information about their order—as handled in short-term memory, and further of making this distinction on the basis of the apparently corresponding kinds of errors. Conrad (1965) has argued against such a distinction, on grounds that are examined elsewhere (Ryan, 1967). It should be noted that such a selective decrease in order errors, in as much as it indicates an improvement in the discrimination of position, is what would be expected if the analogy, drawn in the introduction, with Garner's (1962) "perceptual anchors" holds.

Second, comparing temporal with non-temporal grouping, this experiment appears to have shown that the former is the more effective means of grouping, and that the latter has only a slight effect. This conclusion might be objected to on the grounds that the particular method used to provide non-temporal grouping, of interleaving brief pips, was not the most efficient non-temporal method that could be found, and therefore not a fair test of any difference between the two means of grouping. One reason why the pips might not be very efficient non-temporal demarcators is that they could well have sounded extraneous to the whole sequence. Another reason is that the locating of the pips themselves within the whole sequence might have caused some difficulty (cf. Ladefoged and

Broadbent, 1960); this however is unlikely since the subjects always knew whereabouts the pips would occur. The combined condition of pips and instructions does not lead to recall that is any better than the condition of instructions alone; this suggests that an external non-temporal demarcation of the groups cannot be any more helpful than a purelv subjective one. If there was any scope for improvement in the subjective non-temporal case, it would be expected that the pips, even if they are not the most efficient external demarcators, would reveal at least some improvement. The most likely reason for there being no such scope is the absense of any time interval between groups. Although it is conceivable that a more efficient method of non-temporal grouping than the present ones could be found, it seems very unlikely, in the light of the present results, that it would be as efficient as temporal grouping.

It therefore appears that temporal grouping improves recall primarily because of the presence of an unfilled interval between the groups, rather than simply because of the division of the sequence. This supports the suggestion made above that the pauses allow more processing of the previous items than could occur in an ungrouped or non-temporally grouped sequence. However the slight, although non-significant, improvement that was found with all forms of non-temporal grouping cannot be ignored, nor can the other signs of grouping accompanying this (the changes in the serial position curves and in the relative frequencies of transpositions). It appears that the purely divisive non-temporal aspect of grouping can have some effect, but that a certain time interval is necessary to exploit this more fully. The question then arises, given the superiority of the pauses as a means of grouping, whether the pauses would be equally valuable anywhere within the sequence or whether they have to occur in a particular pattern. If, as seems likely, the divisive aspect of grouping is at all important, then it would be expected that some patterns of division would be more helpful than others. The next experiment was performed to assess the effectiveness of different patterns of temporal grouping.

Experiment II

Procedure

The same length of sequence as before—nine digits—was used, and each sequence was divided into three groups. The number of groups per sequence never varied, but the size of the groups did, between one item and seven. Every combination and permutation of these group sizes was used to give a variety of irregular groupings, with different degrees of disparity between the group sizes, as well as one case of equally sized groups.

Sequences

The rates of presentation and the pause lengths were the same as in the first experiment. The ungrouped condition, UG, was that used above, consisting of the 56 sequences already described. The experimental condition of irregular grouping, IG, consisted of the same number of sequences. This was made up of two examples of each of the following groupings, where the numbers represent the number of items in each group: the combinations 117, 162, 135, 225, 144, 234, and all the permutations of each of these, and also the combination 333. Four different sets of the 56 irregular sequences were compiled, in order to vary both the particular digits used in any serial position, and the order of occurrence of the different groupings. In compiling each of these four sets the digits were chosen with the same constraints as before. The sequences were further checked to ensure that no particular digit occurred disproportionately often in the smaller groups of one or two, where it would have been rather obvious; this checking only resulted in two changes. The order of the different irregular groupings within each of the four sets was determined from random number tables.

Subjects and instructions

Twenty subjects (different from any above) were used for condition IG, so that each particular grouping was tested 40 times—twice on each subject. Five subjects were randomly assigned to each of the four sets of sequences. The instructions, about the length and composition of the sequences, about recalling in order and about not omitting, were all as above. The subjects were further told that there would always be three groups in each sequence, separated by pauses, and that these groups would vary in size between one and seven; they were also told that the different groupings would occur in a random order so that they could not predict which would come next. Some examples of the different groupings were played to them and there were five practice sequences. The experimental session lasted about half an hour and there were two short rests during it.

Results

(i) *Overall performance.* Combining all the different groupings in condition IG (including 333) the mean error score per sequence and per subject was 1·71. The comparable score for condition UG was 2·11. No significant difference was found between these two conditions, using a Mann-Whitney U-test.

(ii) *The different groupings.* The error scores in condition IG varied considerably between groupings. The total scores are shown in Table III. Before combining the errors it was ascertained that there was no significant difference between the scores for the four sets of sequences. It can be seen that the combination 333 gave the fewest errors, and that sequences with groups of six and seven in them the most. A significant overall effect of the different groupings was found,

TABLE III

Total number of all errors for each kind of grouping, combining the scores of all subjects ($N = 20$) *on the two examples of each grouping*

Grouping	Total errors	Grouping	Total errors
333	20	234	64
243	43	414	66
441	45	225	68
144	47	135	69
522	47	324	71
153	51	612	78
513	51	162	86
252	52	117	87
342	54	621	89
351	56	126	95
315	59	261	98
531	60	216	104
423	62	711	108
432	62	171	116

using a Friedman two-way analysis of variance on the individual total scores for each kind of grouping ($\chi^2 = 219{\cdot}4$, $d.f. = 27$, $P < 0{\cdot}001$). Apart from the greater difficulty of sequences containing the larger sized groups, there does not appear to be any consistent pattern in the order of difficulty of the different groupings. Averaging over the permutations of each combination of group sizes did not show any other trend.

To assess which of the different groupings differed significantly from no grouping, the individual mean scores for each kind of grouping were compared to those for condition UG, using a U-test (two-tailed). Only the following groupings showed a significant difference: 333, 423, 432 and 243 ($U = 49$, $P < 0{\cdot}05$; $U = 49$, $P < 0{\cdot}05$; $U = 47$, $P = 0{\cdot}02$; and $U = 52$, $P < 0{\cdot}05$, respectively).

The difference between regular grouping—333—and the apparently easiest irregular grouping (as regards total errors)—243—was found to be significant, using a Wilcoxon test ($T = 20$, $N = 17$, $P < 0{\cdot}01$).

(iii) *Serial position curves.* Percentage serial position curves were constructed for each kind of grouping separately, using the total error scores. In most cases, the resulting curves corresponded visually to the pattern of grouping in question, in that dips in each curve occurred at the positions representing the ends of groups. Some examples of this are shown in Figure 2. An assessment was made of how often the grouped curves fell or remained level at the end of groups, where the ungrouped curve rose instead. Discounting the first and last serial positions, and also the smallest groups of one and two items, but combining all the groupings, the grouped curve fell or remained level in 34 out of 43 positions where the ungrouped one rose.

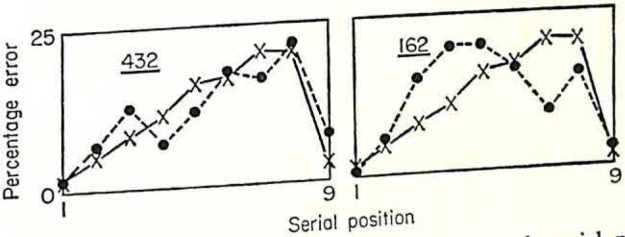


FIG. 2. Curves showing the percentage of the total errors at each serial position:—X—X—, ungrouped;-●- - -●-, irregularly grouped.

Discussion

These results indicate that the effectiveness of temporal grouping in improving recall varies with the pattern of the grouping. No case of irregular grouping was found to be as helpful as regular grouping, and sequences containing the largest groups of six and seven items were recalled slightly worse (although this was not found to be significant) than ungrouped sequences were. However, nearly all the irregular groupings gave rise to a scalloped serial position curve, irrespective of whether there was any overall facilitation or not. This suggests that the pauses can cause local improvements at least, wherever they occur, and even in the face of slightly greater overall difficulty.

It thus appears that the provision of pauses anywhere within a sequence does not necessarily improve recall of the whole sequence; instead, it is important that the pauses should divide up the sequence in certain ways, the most effective pattern of division being a regular one. It might be objected that the results should not be interpreted as a function of the regularity of the groupings, but instead as a function of the particular size of the groups, in this case three. Only a further experiment, using different sequence lengths and different group sizes, could settle this conclusively, but it is unlikely that such an account, apart from its general weakness, could entirely explain even the present results. For example, if size alone were important, it would be expected that irregular groupings containing one group of three (e.g. 135, 234) would be recalled much better than those that do not contain such a group (e.g. 144, 225). This was not found. Further, it would be difficult to explain on grounds of size alone why sequences with groups of six and seven items tended to be recalled worse than ungrouped ones.

While the first experiment emphasized the importance of temporal as opposed to non-temporal grouping, the present experiment has shown how specific the effectiveness of temporal grouping is to a particular pattern of grouping. It is not surprising in an auditory short-term memory task that brief unfilled intervals should be of considerable importance in determining how much is recalled. There is some further evidence, apart from the considerations advanced earlier, about the importance of the time available within a sequence for processing. Aaronson (1967) compared the immediate recall of "compressed" digits with that of normally spoken ones. In compressing the digits the duration of each digit was shortened by 35 per cent, but the duration of each inter-digit interval was correspondingly lengthened, so that the overall duration of the sequence was not altered. The effect of this compression was to improve recall, suggesting that the time intervals between the stimuli of a sequence are particularly important in determining how good subsequent recall is. It was argued above that the precise distribution in time of the incoming items of a sequence might determine how much, or what kind of, processing of previous items is carried out. In this respect it is of interest that the improving effect of compression in Aaronson's experiment consisted entirely in a decrease in order errors, and not in "item" errors; this resembles the improvement found here with temporal grouping. Together these results suggest that the interstimulus intervals are used for the coding of order information.

An explanation of why irregularly distributed pauses are not so helpful as regular ones may lie in this kind of recoding that regular grouping seems to cause, namely an improvement in the coding of order information. In the discrimination of position along a continuum it will in general be advantageous if the continuum can be divided, for the reasons mentioned earlier in connection with Garner's (1962) suggestions about the functions of "perceptual anchors." It can be shown arithmetically (Ryan, 1967) that if the task involves guessing the position of items, equal division of the continuum will result in fewer guesses on average than irregular division into the same number of sections. Although this analogy fits the present results on irregular compared to regular grouping, it can also be shown that sequences of eight items divided into four groups of two should be recalled as well as those divided into two groups of four; this is not found (Wickelgren, 1964). A further explanation of why irregular grouping is inferior to regular grouping may lie in the precise kind of recoding of order information that the latter allows. Some indication of this was obtained from the transposition data of the first experiment, where transpositions one- and two-apart were reduced proportionately more than those three-apart. This kind of recoding might depend on the existence of exactly corresponding positions in different groups (e.g. beginning, middle and end in all groups). Such an exact correspondence does not exist in irregularly grouped sequences.

In general, these experiments on grouping point to the importance of considering what happens during the presentation of a sequence. Small alterations in the time of arrival of items within a sequence can cause considerable changes in the amount recalled, and in the way that information is coded.

I should like to thank Mr. A. J. Watson for his very helpful advice and criticism. I am grateful to Professor O. L. Zangwill for the provision of facilities, to the S.R.C. for a Research Studentship, and to Dr. M. P. M. Richards for reading a draft of this paper.

References

Aaronson, D. R. (1967). Immediate recall of normal and "compressed" auditory sequences. Paper read at Boston meeting of *Eastern Psychol. Assoc.*

Adams, H. F. (1915). A note on the effect of rhythm on memory. *Psychol. Rev.* **22,** 289-99.

Conrad, R. (1959). Errors of immediate memory. *Br. J. Psychol.* **50,** 349-59.

Conrad, R. (1965). Order errors in immediate recall of sequences. *J. verb. Learn. verb. Behav.* **4,** 161-9.

Conrad, R. and Hille, B. (1957). Memory for long telephone numbers. *Post Office Telecomm. J.* **10,** 37-9.

Garner, W. R. (1962). *Uncertainty and Structure as Psychological Concepts.* New York: John Wiley and Sons.

Ladefoged, P. and Broadbent, D. E. (1960). Perception of sequence in auditory events. *Q. Jl exp. Psychol.* **12,** 162-70.

Maynzer, M. S., Tresselt, M. E., Adler, S., Cohen, A. and Schoenberg, K. M. (1966). Short-term retention of digits: a function of item distribution with respect to time. *Psychon. Sci.* **5,** 385-6.

Michon, J. A. (1964). Temporal structure of letter groups and span of perception. *Q. Jl exp. Psychol.* **16,** 232-40.

Miller, G. A. (1956). The magical number seven, plus or minus two: some limits on our capacity for processing information. *Psychol. Rev.* **63,** 81-97.

Murray, D. J. (1965). Vocalization-at-presentation and immediate recall, with varying presentation rates. *Q. Jl exp. Psychol.* **17,** 47-56.

Pollack, I., Johnson, L. D. and Knaff, R. P. (1959). Running memory span. *J. exp. Psychol.* **57,** 137-46.

Ryan, J. F. (1967). *Grouping and Short-term Memory.* Ph.D. Dissertation. University of Cambridge.

Siegel, S. (1956). *Non-parametric Statistics for the Behavioural Sciences.* New York: McGraw-Hill Book Co. Inc.

Wickelgren, W. A. (1964). Size of rehearsal group and short-term memory. *J. exp. Psychol.* **68,** 413-9.

Received 9 *November* 1968

Q. Jl exp. Psychol. (1969) **21**, 148–155

TEMPORAL GROUPING, REHEARSAL AND SHORT-TERM MEMORY

JOANNA RYAN†
Psychological Laboratory, Cambridge

The effects of temporal grouping on auditory short-term memory are investigated, using pauses of different lengths to indicate the groups. No quantitative or qualitative differences were found between the shorter and longer pauses. This result, in conjunction with an analysis of the order errors, is used to suggest what kind of processing occurs during the pauses. This processing is selective, recoding the information about order according to position within a group.

Introduction

In a previous experiment (Ryan, 1969) the effect of grouping in improving auditory short-term memory was investigated; it was found that regular temporal grouping, with the groups demarcated by slight pauses, lead to significantly better recall than did similar non-temporal grouping, with the groups demarcated by brief pips. Non-temporal grouping did not improve recall significantly when compared with an ungrouped condition. It was suggested that the superiority of temporal grouping was due to the subject using the unfilled intervals between groups to process the previous items, without having at the same time to attend to new incoming items. The total time available in an ungrouped or non-temporally grouped sequence for such processing is in principle equal to that available in a temporally grouped sequence, since the sequence durations are not altered. However, it is possible that this time is of more use when several items can be processed at once, than when the subject has continually to alternate one item at a time between these two tasks of attending and processing.

Indirect evidence that the pattern of processing is changed by a grouped presentation has been provided by Kahneman, Onuska and Wolman (1968). They measured pupil size during the presentation of ungrouped and grouped sequences which had to be recalled immediately. With ungrouped sequences the pupil dilated steadily from beginning to end of the sequence; with grouped sequences the pupil size increased much less during the presentation of the whole sequence, but showed considerable variation, dilating slightly at the beginning of each pause, and constricting towards the end. Kahneman *et al.* (1968) tentatively ascribe pupil dilation in such tasks to increased effort in rehearsing previous items; in the ungrouped condition the subject would be trying to rehearse throughout each sequence, in the grouped one, only during the pauses.

---

† Present address: Unit for Research on Medical Applications of Psychology, Station Road, Cambridge.

The present experiment is an attempt to provide more information about the kind of processing that can occur in a temporally grouped sequence but not in an ungrouped one. In particular it is concerned with the possibility that grouped sequences are recalled better simply because more processing can occur than with ungrouped sequences. Two grouped conditions are used, in one of which the length of the pauses between groups is increased. Lengthening the pauses should allow time for more processing to occur; if the processing is such that its quantity is important, then lengthening the pauses should be additionally beneficial when compared with the equivalent ungrouped condition. If the quantity of processing is not important, and no additional improvement is obtained, qualitative changes in the pattern of recall errors might instead provide some information about differences in processing between grouped and ungrouped sequences. One reason for supposing that the quantity of processing will be important derives from experiments on post-presentation rehearsal intervals. For example, Peterson and Peterson (1959) found that the longer the interval during which the subjects rehearsed aloud before the onset of an interfering activity, the better the subsequent recall. Other experiments with sequences of items (Murdock, 1961; Sanders, 1961; Pollack, 1963), although not using explicit instructions to rehearse found similar results. These experiments all suggest that the longer there is to rehearse previous items, the more resistant they become to any interference. If the processing that occurs in temporally grouped sequences between the groups resembles this kind of rehearsal, then it would be expected that lengthening the pauses should lead to an additional improvement in recall.

Method

There were four conditions, an ungrouped and a grouped condition for both of two different sequence durations. In lengthening the pauses between groups, the total duration of a sequence is also lengthened, and this increase in total duration might have effects independent of the increase in pause length. For example, longer duration might lead to worse recall, on grounds of decay, and this could obscure any beneficial effects of more processing. In order to control for such confounding effects two ungrouped conditions were used, so that each grouped condition (short and long pauses) could be compared with an ungrouped one of the same total duration. Thus the relative effects of the two grouping conditions could be assessed free of any confounding factors, as well as any absolute difference between them.

Sequences

Nine-item sequences, and grouping into three groups of three were used, as this had previously been found to give convenient levels of difficulty. Letters were used to compose the sequences; these were chosen in preference to digits as they are harder to recall, and therefore allow more scope for improvement than do digits. Since one possible outcome of the experiment was better recall with the longer pauses it was important to allow this to show. A vocabulary of 10 letters was used, chosen so as to minimize the possibility of acoustic confusion. These were F H J K L N P R W Z. The sequences were compiled with the use of random number tables and the following constraints: there were no repetitions of any one letter within a sequence, and each letter occurred with approximately equal frequency in each serial position, over all the sequences. Forty sequences were used for the ungrouped conditions, 20 at each rate of presentation. The same 40 sequences were used for the grouped conditions, 20 for each pause length.

Rate of presentation

The rate of presentation for the faster ungrouped condition, UF, was one item every 0·55 sec., each sequence lasting for nearly 5 sec. In the corresponding grouped condition, GF, the total duration of each sequence was the same as in condition UF. Within each group the rate of presentation was one item every 0·45 sec., giving a pause length (from the beginning of the last item in one group to the beginning of the first item in the next group) of 0·9 sec. In the slower ungrouped condition, US, the rate of presentation was one item every 1·1 sec., thus doubling the sequence duration. In the corresponding grouped condition, GS, the sequence duration was the same as in condition US. The rate of presentation within groups was the same as in condition GF, namely one item every 0·45 sec., and the length of the pauses between groups was 3·4 sec. The interval between sequences was 10 sec.

Subjects and design

Forty subjects were used altogether, drawn from the undergraduate population of Cambridge University. 20 of these were tested on both of the two ungrouped conditions, and the remaining 20 on both of the two grouped conditions. The design was counter-balanced so that within each of these conditions half the subjects took a fast rate first, and half a slow one first. The assignment of subjects to conditions and order of rates was random.

Instructions

The subjects were told that all the sequences were nine items in length, and that they were composed of randomly arranged letters drawn from a vocabulary of ten, with no repetitions within any one sequence. They were presented with a list of the ten possible letters, which they were required to read through twice and keep in front of them during the experiment. They were told to recall aloud as quickly as possible, in the correct order, and to guess if they could not remember. It had previously been found that such instructions not to omit were followed successfully. The experimenter wrote down what the subjects said. The subjects were also told whether to expect ungrouped or grouped sequences, and in the latter case they were told that the grouped sequences contained three groups of three, separated by pauses. They were further told that two different rates of presentation would be used, and in what order to expect them. Five practice sequences at the appropriate rate were given before each set of 20 sequences. Subjects were tested individually and the whole experiment lasted about half an hour.

Scoring

The sequences were scored by counting the individual errors in each sequence, for each subject. The errors were categorized according to the loss they represented from each serial position, as follows: (i) transposition—any item not in its correct serial position, but occurring elsewhere in the sequence; (ii) item error—any item neither in its correct position, nor included elsewhere in the sequence.

Results

The scores for all subjects and all errors combined are shown in Table I, for each condition separately.

Effect of different rates

The scores were first examined to see if there were any transfer effects between rates within each condition (ungrouped and grouped), which would prohibit combining the scores of subjects tested in a different order on the same rate. Taking each rate separately, the scores of the ten subjects taking one rate first

TABLE I

Total number of all errors, for both conditions and rates of presentation

Condition and rate	Number of errors Tested first ($N = 10$)	Tested second ($N = 10$)	Total ($N = 20$)
Ungrouped			
Fast UF	883	864	1747
Slow US	884	904	1788
Grouped			
Fast GF	715	682	1397
Slow GS	656	697	1353

were compared with the scores of the 10 subjects taking it second, using a two-tailed Mann-Whitney U-test. For both ungrouped and grouped conditions there was no significant effect of the order in which the rates were tested (values of $P \leqslant 0{\cdot}05$ being required). This was true for both rates. Within each rate, therefore, the scores of all 20 subjects could be combined, irrespective of the order of testing.

The difference between the fast and slow rates was then assessed, using a two-tailed Wilcoxon matched-pairs test; for neither ungrouped nor grouped condition was any significant difference found between the rates.

Effect of grouping

For each rate the improvement with grouping was assessed with a one-tailed Mann-Whitney U-test. For both rates the improvement was significant (fast, $U = 131$, $P < 0{\cdot}05$; slow, $U = 117$, $P < 0{\cdot}02$). For the fast rate the mean percentage improvement with grouping was 21·5 per cent; for the slow, it was 24·7 per cent.

Error analysis

The number of different kinds of errors—transposition and item errors—is shown in Table II, for both rates and conditions.

It should be noted that the improvement with grouping is almost entirely due to a reduction in the transposition errors. There was no significant difference between the two rates in the number of different errors, for either condition. Analysis of the transposition errors only, into those adjacent, two-apart and three-apart showed no effect of rate. There was however an effect of grouping, for both rates, in that the adjacent and two-apart transpositions were reduced considerably by grouping, whereas the three-apart ones increased. In both the ungrouped conditions there is an inverse relationship between the frequency and distance of transpositions. In the grouped conditions this relationship no longer

TABLE II

Total number of item and transposition errors for each condition and rate of presentation ($N = 20$)

Condition and rate	Number of errors Item errors	Trsps. total	Trsps. adjacent	Trsps. 2-apart	Trsps. 3-apart	Trsps. others
Ungrouped						
Fast UF	787	960	484	242	141	90
Slow US	771	1017	479	244	162	132
Grouped						
Fast GF	805	592	208	111	197	76
Slow GS	763	580	180	112	199	89

exists; instead the three-apart errors are more numerous than the two-apart ones and are roughly equal to the adjacent ones. The significance of this change in the transposition pattern was assessed in terms of the number of subjects in each condition who individually conformed to either of the overall transposition patterns described. For both the rates, the changes between ungrouped and grouped conditions were significant (fast, $\chi^2 = 18\cdot4$, $d.f. = 2$, $P<0\cdot001$; slow $\chi^2 = 17\cdot0$, $d.f. = 2$, $P<0\cdot001$).

The adjacent and two-apart transpositions were analysed further, into those occurring between different groups (e.g. third position in one group to second in next), and those occurring within the same group (e.g. third position to second in the same group). For the purpose of the analysis, shown in Table III, the ungrouped sequences were divided into comparable groups of three. It can be seen that both transpositions within groups and those between are reduced by grouping for both rates; however proportionately the adjacent errors within groups are reduced less than the adjacent errors between groups. The reverse is

TABLE III

Transpositions within and between groups, for all subjects ($N = 20$)

Condition	Within		Between	
	Adjacent	2-apart	Adjacent	2-apart
Ungrouped				
Fast UF	368	71	119	171
Slow US	373	69	106	175
Grouped				
Fast GF	182	26	26	85
Slow GS	166	25	14	87

true of the two-apart errors, those within groups being reduced proportionately more than those between.

Discussion

No difference was found between temporal grouping with pauses of 0·9 sec., and with pauses of 3·4 sec. It thus appears that lengthening the pauses between groups does not enhance the effects of temporal grouping in any way, despite the superiority of temporal grouping as compared to non-temporal grouping. It was suggested earlier that the amount of processing occurring in the pauses could be an important factor in improving recall. This suggestion is not tenable in the light of the present results, since the opportunity that the lengthened pauses provide for additional processing is not of any use. This finding differs from the results of the experiments cited above on post-presentation rehearsal, where longer rehearsal intervals were found to lead to progressively better recall.

It might be objected that any improving effect of lengthening the pauses would be masked by an adverse effect of lengthening the total duration of each sequence. Such an adverse effect would be expected according to a decay theory, and has been found in some experiments on rate of presentation (e.g. Conrad and Hille, 1958; Posner, 1964); others however (e.g. Pollack, 1952; Aaronson, 1968) have found an opposite beneficial effect. These, and other conflicting findings cannot be entirely accounted for by the use of different rates; the present experiment adds to the confusion by showing no difference in the ungrouped condition between the two rates used. This may be because the range used here (one item every 0·55 sec., and one every 1·1 sec.) is not large enough to show any difference, being smaller than the range used in most other experiments. Alternatively, this particular range may be the least sensitive one; the possible consequences of using different rates are too various for it to be likely that the effects of changing rates will be the same for all rates. Whatever the explanation, it appears that the effects of total duration are not important here, in the ungrouped condition. It is a reasonable, although disputable, inference that such effects are not therefore operative in the grouped condition, and thus are not likely to mask any other independent effect of lengthening the pauses.

It might further be objected that the increase in the size of the pauses was not sufficient to show any effect, whereas a further increase might. This is unlikely to be true, since the length of the pauses was more than trebled; in the previous experiment (Ryan, 1969) comparing non-temporal and temporal grouping the interval between the groups was only increased from 0·55 to 0·9 sec., that is, not even doubled, and a significant improvement was found.

Assuming then that the present experiment provides a valid means of testing whether lengthening the pauses enhances the effects of temporal grouping, the question of what kind of processing occurs in the pauses still remains to be answered. Some clues may be obtained by considering the changes in the error patterns, which are the same for both pause lengths. The function of the processing appears to be specific rather than general, since it is almost exclusively the transposition errors that benefit from the grouping. This confirms previous findings (Ryan, 1968). As well as this selective improvement in the recall of the order of items, there appears to be a change in the way in which such information is coded. This is seen in the relative frequencies of the different transpositions, when analysed for distance. The inverse relationship between distance and frequency found in the ungrouped condition is replaced by one in which the adjacent transpositions are about equal to the three-apart ones, and the latter are more numerous than the two-apart ones. This is at first sight curious, in that transpositions between different groups to the same relative position within each group are not reduced at all by grouping, whereas all other transpositions are. Since the relative number of errors, and thus the confusion between different groups is increased in this way, it is unlikely that the recoding principle is supplied primarily by which group an item is in. That is, it is unlikely that the items of one group are coded together as a unit or "chunk," separate from the items of other groups. If this were so fewer transpositions between groups would be expected, and possibly more within groups, compared to the ungrouped condition. The changes in the adjacent and two-apart errors complicate matters further. Both those between different groups and those within groups are reduced by grouping. Proportionately the adjacent errors between groups are reduced more than those within groups; the reverse is true for the two-apart errors, those between groups being reduced proportionately less than those within. These results may be at least partially explained if it is supposed that grouping recodes the items by labelling their positions repetitively, according to the position within each group, thus: 123123123, instead of from 1 to 9, as in an ungrouped sequence. There then appears to be a direct relationship in the extent to which the different transpositions are reduced by grouping and the extent of the numerical change in label involved. Grouping reduces most those transpositions which involve the greatest change in label (i.e. 1 to 3) regardless of whether these concern adjacent or two-apart errors, or are between or within groups. Those errors which involve only a change of one in their label (e.g. 1 to 2, 3 to 2) are reduced to a lesser extent, again irrespective of whether these are adjacent or two-apart, and of whether they are between groups or within. Finally those transpositions which involve no change in label (e.g. 1 to 1) are not reduced at all. If the suggested system of recoding or labelling operates by making the positions of items more distinct in this way, then this kind of relationship in the relative confusability of positions would follow.

If it is supposed that the pauses in a temporally grouped sequence are used to execute such recoding, it still needs to be explained why lengthening the pauses does not improve recall additionally. It might be argued that having more time in which to recode the items should improve the recoding. One somewhat ad hoc explanation is that the recoding has to take place quickly before the unprocessed items decay, and that once done, the recoding cannot be revised in any way, the information on which it was based having been discarded in favour of the recoded version.

This experiment has provided some evidence about the kind of processing that occurs in a temporally grouped sequence, and that is associated with improved recall. Should this processing be called rehearsal? "Rehearsal" is a term that is frequently used, but seldom precisely defined. The reasons for this probably lie both in its normally covert nature, and in the very general preserving functions that are ascribed to it. If the processing referred to above is identified with rehearsal—and there do not seem to be any compelling reasons against doing this—then it appears that rehearsal can operate selectively, on only part of the total information load. It further appears that not all rehearsal is the same, in that the rehearsal presumed to occur during the presentation of a sequence is not such that its quantity is of great importance, whereas the rehearsal presumed to occur after the presentation of a sequence and before any interference is such that its quantity influences subsequent recall.

I should like to thank Mr. A. J. Watson for his invaluable help and criticism. I am grateful to Professor O. L. Zangwill for providing facilities, to the S.R.C. for a Research Studentship, and to Dr. M. P. M. Richards for reading a draft of this paper.

References

AARONSON, D. R. (1968). The temporal course of perception in an immediate recall task. *J. exp. Psychol.* **76**, 129–40.

CONRAD, R. and HILLE, B. (1958). Decay theory of immediate memory and paced recall. *Canad. J. Psychol.* **12**, 1–7.

KAHNEMAN, D., ONUSKA, L. and WOLMAN, R. (1968). Effects of grouping on the pupillary response in a short-term memory task. *Q. Jl exp. Psychol.* 20, 309–11.

MURDOCK, B. B. (1961). The retention of individual items. *J. exp. Psychol.* **62**, 618–25.

PETERSON, L. R. and PETERSON, M. (1959). Short-term retention of individual items. *J. exp. Psychol.* **58**, 193–8.

POLLACK, I. (1952). The effect of rate of presentation of information. H.F.O.R.L. Rpt. No. 25; AD 140.

POLLACK, I. (1963). Interference, rehearsal and short-term retention of digits. *Can. J. Psychol.* **17**, 380–92.

POSNER, M. (1964). Rate of presentation and order of recall in immediate memory. *Br. J. Psychol.* **55**, 303–7.

RYAN, J. (1969). Grouping and short-term memory: different means and patterns of grouping. *Q. Jl exp. Psychol.* **21**, 137–47.

SANDERS, A. F. (1961). Rehearsal and recall in immediate memory. *Ergonomics*, **4**, 25–34.

Received 9 *November* 1968

Q. Jl exp. Psychol. (1969) **21**, 156–161

PARTIAL REINFORCEMENT AND PARTIAL DELAY OF REINFORCEMENT EFFECTS WITH 72-HOUR INTERTRIAL INTERVALS AND INTERPOLATED CONTINUOUS REINFORCEMENT

MICHAEL E. RASHOTTE† AND C. THOMAS SURRIDGE
University of Toronto

Three groups of rats ran 108 trials in a straight runway, one trial every 3 days. On the first 44 trials, one group received continuous (and immediate) reinforcement (CRF), a second group 50 per cent partial reinforcement (PRF), and the third group a 50 per cent schedule of partial delay of reinforcement (PDR). All groups received CRF on the next 20 trials, and extinction on the last 44 trials. The PRF and PDR groups extinguished at approximately the same rate, and significantly more slowly than the CRF group.

Introduction

Resistance to extinction of a running response is increased by randomly omitting reinforcement (PRF) or by delaying reinforcement for a sufficient length of time (PDR) on some percentage of trials (e.g. Crum, Brown and Bitterman, 1951; Surridge, Boehnert and Amsel, 1966; Theios, 1962). Recently, the similarity between PRF and PDR effects has been further emphasized by the demonstration that PDR produces this effect even under the relatively extreme training conditions employed in some PRF experiments: widely spaced (24 hr.) trials with continuous immediate reinforcement (CRF) interpolated between acquisition and extinction (Donin, Surridge and Amsel, 1967).

In the present experiment we investigated this similarity more directly and under even more extreme conditions of trial spacing than previously employed. PRF and PDR groups were trained with 72-hr. intertrial intervals (ITI) and CRF was interpolated before extinction. A CRF group served as a baseline control.

Method

Subjects

Twenty-four male Wistar albino rats, 60 to 65 days old were supplied by Woodlyn Farms, Guelph, Ontario. They were maintained on a 10-g. per day food deprivation regimen and *ad libitum* water, and were handled and tamed for 17 days prior to the start of experimental training.

---

† Present address: Department of Psychology, Florida State University, Tallahassee, Fla. 32306, U.S.A.

Apparatus

The apparatus was a straight enclosed runway, 3¾ in. deep and covered with a clear plexiglas lid, which consisted of a 13¼-in. start box painted flat grey and a 5-ft. 3-in. alley painted flat black. The goal area comprized the last 14 in. of the alley and was separated from the rest of the alley by an aluminium guillotine retrace door. A Model D Gerbrands pellet dispenser provided reinforcement when the final photobeam was interrupted. A Hunter timer in this circuit permitted delay between the breaking of the photobeam and activation of the reinforcement device.

Five 1-ft. running-time measures were provided (to 0·01 sec.) beginning with the opening of the start door and ending 3 in. from the end wall of the goal area.

Procedure

Experimental design. The subjects were randomly assigned to three groups of 8 for 44 acquisition trials: a continuous and immediate reinforcement group (Group CRF), a partial reinforcement group (Group PRF), and a partial delay of reinforcement group (Group PDR). All groups then recived CRF for 20 trials followed by 44 trials of extinction.

Experimental training. Subjects were run in two squads of 12, four subjects from each group in each squad. The order in which subjects were run on each trial was randomized. On reinforced trials the subject was removed from the goal box immediately after eating and on all non-reinforced trials confinement duration was 60 sec. All 12 subjects in a squad were returned to their home cages after the day's trial and given their daily food ration 30 to 40 min. later. The subjects were given one trial every 3 days over a 324-day period. Trials were given at the same time on each running day and subjects were fed their daily ration at the same time on all days.

The maximum time recorded in any measure was 30 sec. In extinction, the subject was removed from the runway and placed in the goal box for the 60-sec. confinement period if it had not broken the final photobeam 180 sec. after the start door was opened. A time of 30 sec. was recorded for the measures which had not been traversed.

Phase 1 consisted of 44 trials during which reinforcement was omitted on half the trials for Group PRF, delayed for 30 sec. on half the trials for Group PDR and given immediately on all trials for Group CRF. When reinforced, subjects received four 250-mg. Noyes pellets (1 g.). Delayed or non-reinforced trials occurred four times in every eight-trial block according to the following sequences in which A indicates an immediately rewarded trial and B a non-reinforced or delayed trial: AABBBAAB, BAABAABB, ABBABBAA, BBAAABBA. Sequences of eight trials were randomly assigned to subjects. On the last four trials of the 44-trial Phase 1, subjects received their delayed or non-reinforced trials according to the first four trials of one of the eight-trial sequences. In phase 2, all groups were run 20 trials to CRF. For subjects in Groups PRF and PDR, but not in Group CRF, this represented a change in reward schedule. Phase 3 consisted of 44 extinction trials for all subjects.

Results

Time scores on each trial for each of the five performance measures were converted to reciprocals (ft./sec.). Figure 1 plots mean speeds in the first (Start), third (Run 2) and fifth (Goal) measures in four-trial blocks for the three experimental phases. The data of the second and fourth measures were omitted from Figure 1 since they were virtually identical with the Run 2 measure.

Analyses of variance performed separately for each measure on speeds over the 11 acquisition blocks yielded reliable block main effects only ($P<0{\cdot}01$ in all measures), indicating that although acquisition had occurred, there were no significant differences in performance among the three groups. Analyses of speeds throughout the interpolated CRF training period also yielded no significant

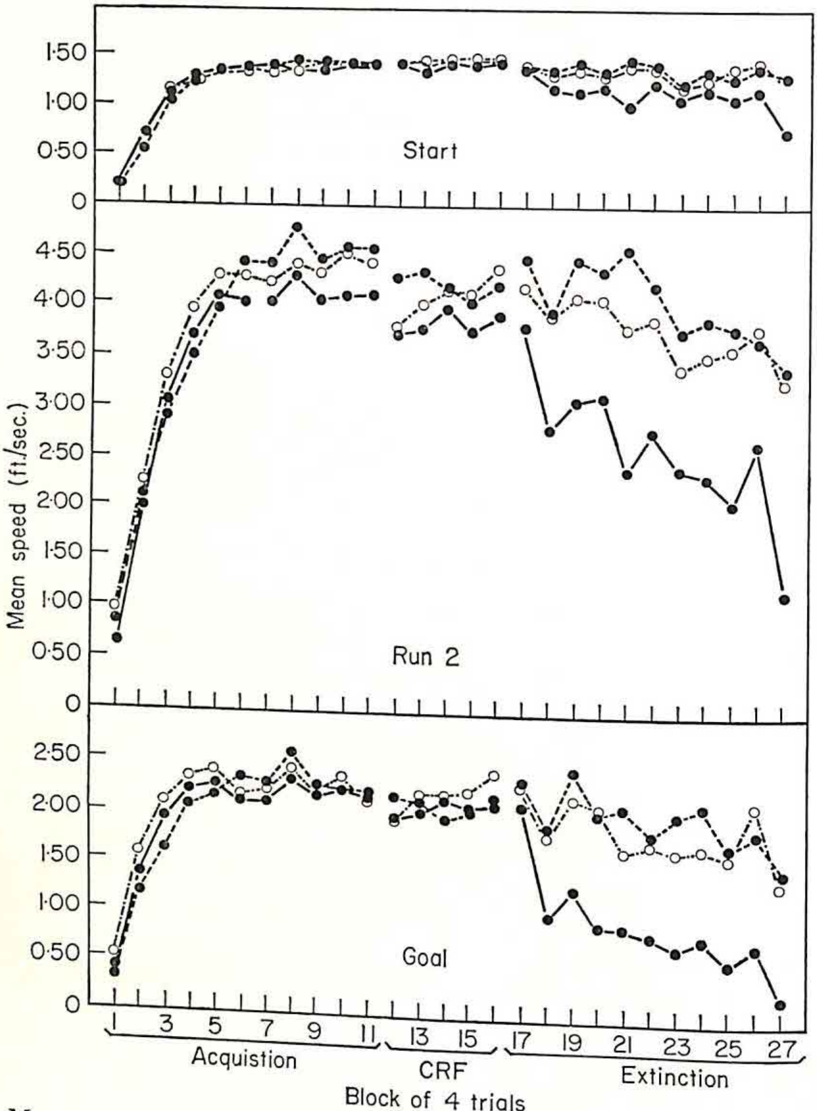


FIG. 1. Mean speeds in the first, third and fifth alley segments throughout acquisition, interpolated CRF training and extinction. ●—●, CRF; ●- - -●, PRF; ○—.—.—○, PDR.

F values for Group or Group × Block factors and, with the exception of an inexplicable drop in speed for all groups in Run 2, there was little difference between terminal CRF training performance and terminal acquisition performance.

Figure 1 suggests that Groups PRF and PDR extinguished less rapidly than Group CRF and, in fact, showed little performance decrement in any measure over the 44 extinction trials.

For purposes of statistical analysis, speed on Block 16 and on each extinction block was expressed as a proportion of speed on Block 16 for each subject. Analyses of variance were performed on these relative speeds comparing performance among the three groups over Block 16 and the 11 extinction blocks. These analyses yielded, for Start, Run 2 and Goal, respectively, reliable Group main effects,

$F(2,21) = 5\cdot23$ ($P<0\cdot05$), $6\cdot39$ and $15\cdot70$ ($P<0\cdot01$); Block main effects, $F(11,231) = 4\cdot94$, $16\cdot76$ and $23\cdot15$ ($P<0\cdot01$); and Group × Block interactions, $F(22,231) = 2\cdot12$, $2\cdot76$ and $3\cdot67$ ($P<0\cdot01$).

Individual comparisons between the group means indicated that in all measures, both PRF and PDR differed significantly from CRF (Newman-Keuls $P<0\cdot05$) but not from each other. In order to isolate the source of the Group × Block interactions, the sums of squares for each measure were partitioned into the PDR vs. PRF × Block, and the PDR-PRF vs. CRF × Block components. Groups PDR and PRF did not interact significantly over Blocks; $F(11,231) < 1\cdot00$ for Start and Run 2 and $= 1\cdot48$ for Goal ($P<0\cdot05$). However, the PDR-PRF vs. CRF × Block interaction was significant, $F(11,231) = 3\cdot88$, $4\cdot65$ and $7\cdot07$ ($P<0\cdot01$).

The results of these analyses are consistent with the suggestion from Figure 1 that the extinction performance of Groups PDR and PRF did not differ, and that these groups exhibited greater resistance to extinction than Group CRF.

Discussion

No reliable difference was found among the PRF, PDR and CRF groups during acquisition. This finding extends earlier reports that PDR and CRF produce similar acquisition performance under both spaced (24 hr.) (Donin *et al.*, 1967; Wike and Kintsch, 1959, Exp. 2) and highly massed trials (Crum *et al.*, 1951). These experiments and our own all employed long delays of reinforcement (20 to 30 sec.), random scheduling of delayed and immediate reinforcement trials, and similar delay procedures (subjects were delayed in the goal box). However, our failure to reject the null-hypothesis in the analysis of acquisition speeds is inconsistent with Goodrich's (1959) finding that PRF speeds are higher than CRF in the early portions of the response chain and lower in the later portions. Although the Goodrich effect has been found in experiments which employ 24 ITIs (e.g. Surridge *et al.*, 1966), it does not occur with regularity in between-group partial reinforcement experiments (e.g., Amsel, Rashotte and MacKinnon, 1966, Exp. 3 and 4; Rashotte and Amsel, 1968) and the variables controlling its occurrence have not been specified completely. The present data constitute another instance of failure to obtain the effect.

The extinction data confirm the results of Donin *et al.*, (1967) that PDR, with 30-sec. delays, increases resistance to extinction even after interpolated CRF. They also support Theios' (1962) finding that the PRF extinction effect carries through interpolated CRF. Both of these previous findings are extended by the present data to training conditions in which trials are separated by 72 hr.

Our extinction data do not support the prediction from Capaldi's sequential analysis of PRF and PDR that resistance to extinction is greater after PRF than after PDR if sequences of reinforcements and non-reinforcements are the same as sequences of non-delays and delays (Capaldi, 1967, p. 97). Other direct comparisons between PRF and PDR, in which either non-reinforcement or 20-sec. delay of reinforcement regularly alternated with immediate reinforcement throughout acquisition, have yielded conflicting results: Cogan and Capaldi (1961) found more resistance to extinction for the PDR group whereas Burt and Wike (1963) found more for the PRF group. Capaldi and Olivier (1966) found that PRF led to more resistance to extinction than PDR when the sequence of 15-sec. goal box confinements on non-reinforced trials was identical to the sequence of 10-sec. reinforcement delays. Of course, Capaldi's prediction is dependent on the assumption that delay of reinforcement produces a stimulus after-effect which differs from that produced by non-reinforcement and this assumption requires appropriate values of several parameters of these goal events, such as length of delay, confinement duration on non-reinforced trials, and magnitude of reinforcement. In the present experiment we failed to find a difference between PDR and PRF when the length of delay was 30 sec., the non-reinforcement confinement duration was 60 sec. and the magnitude of reinforcement was 1 g.

This experiment is one of a series which has provided evidence for the similarity of PDR and PRF (Donin *et al.*, 1967; Surridge, Mock and Amsel, 1968). While our present results do not demand it, they are compatible with the proposal made in these earlier experiments that this similarity is due to frustrative reactions evoked by both PRF and PDR which, in classically conditioned form, modulate instrumental responding (Amsel, 1958).

This work was supported by research Grants GB 3772 from the National Science Foundation and APB-72 from the National Research Council of Canada awarded to Abram Amsel.

References

AMSEL, A. (1958). The role of frustrative non-reward in non-continuous reward situations. *Psychol. Bull.* **55,** 102–9.

AMSEL, A., RASHOTTE, M. E. and MACKINNON, J. R. (1966). Partial reinforcement effects within-subject and between-subjects. *Psychol. Monogr.* **80,** (20, Whole No. 628).

BURT, D. H. and WIKE, E. L. (1963). Effects of alternating partial reinforcement and alternating delay of reinforcement on a runway response. *Psychol. Rep.* **13,** 439–42.

CAPALDI, E. J. (1967). A sequential hypothesis of instrumental learning. In SPENCE, K. W. and SPENCE, J. T. (Eds.), *The Psychology of Learning and Motivation.* Vol. I (*Advances in Research and Theory*). New York: Academic Press.

CAPALDI, E. J. and OLIVIER, W. P. (1966). Sequence of delayed reward and non-rewarded trials. *J. exp. Psychol.* **72,** 307–10.

COGAN, D. and CAPALDI, E. J. (1961). Relative effects of delayed reinforcement and partial reinforcement on acquisition and extinction. *Psychol. Rep.* **9,** 7–13.

CRUM, J., BROWN, W. L. and BITTERMAN, M. E. (1951). The effect of partial and delayed reinforcement on resistance to extinction. *Am. J. Psychol.* **64,** 228–37.

DONIN, J. A., SURRIDGE, C. T. and AMSEL, A. (1967). Extinction following partial delay of reward with immediate continuous reward interpolated, at 24-hr. intertrial intervals. *J. exp. Psychol.* **74,** 50–3.

GOODRICH, K. P. (1959). Performance in different segments of an instrumental response chain as a function of reinforcement schedule. *J. exp. Psychol.* **57,** 57–63.

RASHOTTE, M. E. and AMSEL, A. (1968). The generalized PRE: within-*S* PRF and CRF training in different runways, at different times of day, by different experimenters. *Psychon. Sci.* **11,** 315–6.

SURRIDGE, C. T., BOEHNERT, J. B. and AMSEL, A. (1966). Effect of interpolated extinction on the reacquisition of partially and continuously rewarded responses. *J. exp. Psychol.* **72,** 564–70.

SURRIDGE, C. T., MOCK, K. R. and AMSEL, A. (1968). Effect of interpolated extinction after partial delay of reward training on subsequent reacquisition and extinction. *Q. Jl exp. Psychol.* **20,** 321–28.

THEIOS, J. (1962). The partial reinforcement effect sustained through blocks of continuous reinforcement. *J. exp. Psychol.* **64,** 1–6.

WIKE, E. L. and KINTSCH, W. (1959). Delayed reinforcement and runway performance. *Psychol. Rec.* **9,** 19–28.

Received 28 *January* 1969

Q. Jl exp. Psychol. (1969) **21**, 162–171

THE EFFECT OF HUNGER ON WATER INTAKE IN RATS

KEITH OATLEY AND D. A. TONGE

Laboratory of Experimental Psychology, University of Sussex, Brighton

Although reciprocal inhibition between eating and drinking has been postulated, the commonly observed reduction of water intake by hungry rats may not be due to any direct inhibitory mechanism. In one experiment rats deprived of food for 24 hr. and then injected with 2 ml. of NaCl drank the same as rats that had received food *ad lib*. In the second experiment a stomach load of 10 ml. of water 3 hr. before the salt injection was designed to abolish any water deficit that might have occurred during food deprivation had there been inhibition of drinking by hunger. Here again rats deprived of food did not drink less than undeprived animals. In fact the hungry rats drank slightly but significantly more. This phenomenon may be related to the observation confirmed here that during food deprivation rats excrete about twice as much rather dilute urine as during *ad lib* food intake. This seems to indicate that though in general food deprived rats drink less than normal they actually drink more than they need. Schedule induced polydipsia also occurs during food deprivation and this too makes it unlikely that water-intake is actually inhibited during hunger.

Introduction

There is a well-known interaction between hunger and thirst such that animals when deprived of food will reduce their intake of water, and when deprived of water will cut down their food-intake (e.g. Verplanck and Hayes, 1953; Cizek, 1959). Reduction of eating when thirsty seems comprehensible as part of the maintenance of constant osmotic pressure in the body fluids (Adolph, 1947) for any introduction of solutes from the food into the bloodstream of a thirsty animal would tend to raise osmotic concentration. This aspect of hunger-thirst interaction has also been reviewed by Jacobs (1964). It is plausible to postulate (Oatley, 1967) that in this situation an inhibitory neural connection exists from either osmoreceptors or the water-intake controller, to the mechanism that regulates food intake. This conclusion is reinforced by the observation (Oatley, 1965) that rats on a 23-hr. food deprivation schedule will only eat about one-third of their normal intake during the 1 hr. eating period if injected with 2 ml. of 6 per cent sodium chloride and not allowed water. This inhibition is completely abolished by allowing the hypertonically injected rats access to water while eating. Control rats injected with isotonic saline eat the same amount whether or not they have water available. The inhibitory mechanism may well be quite complex since McFarland (1964) has reported experiments on doves which suggest that appetitive aspects of hunger are not depressed as much as food intake when the birds are made thirsty.

As to the effect that inducing hunger has on thirst, Miller (1960) has found that electrical stimulation producing eating also led to cessation of drinking in somewhat thirsty rats, and Grossman (1962) has shown that hypothalamic stimulation with adrenergic transmitter substances diminished thirst at the same time that it potentiated eating. These observations have encouraged some (e.g. Grossman, 1967) to believe that hunger inhibits water intake by means of a central mechanism. There have been strong suggestions however that the depressant effect of hunger on thirst might be accounted for in a way which does not involve hypothalamic inhibition. Bolles (1961) points out that "the animal that has been deprived of food has been relieved of one of its chief needs for water," and McFarland, (1964, 1965) reaches a similar conclusion. The experiments of Lepkovsky, Lyman, Fleming, Nagamo and Dimick (1957) have shown how the ratio of food to water in the gut stays constant, and that the preferred source of some at least of this water is from drinking. They go a long way towards explaining the phenomenon of prandial drinking, and also provide support for the argument that water requirements become less during starvation. Furthermore, Radford (1959) has shown that in the long term as well, renal water losses of rats are determined almost entirely by the quantities of solutes which are eaten in the diet and have to be excreted.

The apparent inhibition of thirst during hunger may thus be mediated entirely peripherally, via changes in the body-fluids, and the following experiments comparing the drinking response to salt injections in hungry and satiated rats were carried out to determine whether the depression of drinking is brought about in this way.

In the first experiment it is argued that if reduction of water intake simply reflects reduction of water requirements then a hungry rat's drinking following a hypertonic saline injection would be the same as that of a non-hungry rat, since each would merely be responding to the same experimentally imposed changes in body-fluid composition and attempting to neutralize the osmotic load. On the other hand, if during hunger there were direct central inhibition of thirst in the sense that the signal representing the real level of water need is multiplied by some factor less than unity, then responses to a salt injection would be less in hungry than in non-hungry rats.

Experiment I

Procedures

Twelve male hooded rats, aged 2½ months with body weight ranging from 185 to 231 g., took part in the experiment. They were housed individually in wire cages and given food (diet 41B) and tap water *ad lib* except where otherwise specified. Drinking tests took place in a rack of compartments each measuring 23 × 15 × 16 cm. and fitted with a burette and drinking spout from which fresh tap water could be obtained. The burettes were graduated in 0·2 ml. divisions and were read to the nearest 0·1 ml. No food was available in the drinking compartments.

Before the experiment proper each rat was given two pre-test drinking trials in the drinking rack. These served to accustom the animals to the procedure and also to allow determination of base-line drinking scores. Each rat was taken, injected with 2 ml. of 1 M NaCl, left in his home cage without food or water for 30 min., and then allowed to drink for 1 hr. in a drinking compartment. These pre-test trials were separated by an interval of 2 days and followed by an interval of 4 days.

The experiment proper began with half of the rats being deprived of food for 24 hr. Thirty minutes before the expiry of the 24 hr. period animals were taken at 1 min. intervals. injected intraperitoneally with 2 ml. of 1 M NaCl, left to wait without food or water for 30 min., and then placed in the drinking rack for 1 hr. (Due to the limited capacity of the drinking rack the rats were treated in two batches, each of 3 hungry and 3 non-hungry rats. The second batch started $1\frac{3}{4}$ hr. after the first.)

The second part of this experiment, designed to illustrate the familiar pattern of lesser drinking in hungry as compared with non-hungry rats, followed after an interval of 1 week. A similar procedure was adopted, except that instead of making the animals thirsty by administration of salt, they were deprived of water for 24 hr. before the drinking test. Thus 6 rats were deprived of both food and water, and 6 of water only.

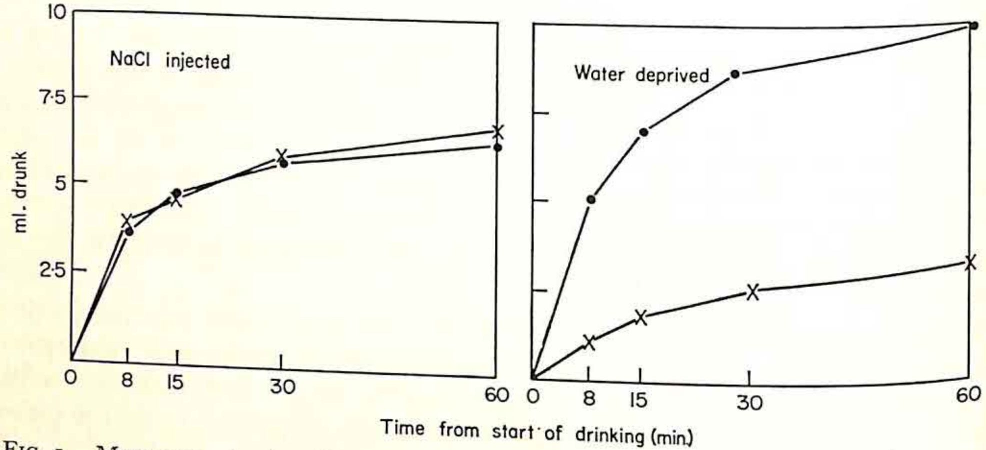


FIG. 1. Mean water intake of hungry and non-hungry rats in Experiment I in response to a hypertonic load (left-hand graph) and water deprivation (right-hand graph). (X) Deprived of food; (●) not deprived of food.

Results and discussion

The results of the drinking tests may be seen in Figure 1. The mean differences between the rats which had been deprived of food and those which had not, were insignificant for drinking in response to a sodium chloride injection (left-hand graph). Respectively *t* was found to be 0·43, 0·15, 0·18 and 0·56 after 8, 15, 30 and 60 min. of drinking. In every case $P > 0{\cdot}5$ with 10 *d.f.* on two-tailed tests.

By comparison, when drinking was induced by deprivation of water, suppression of drinking in the hungry rats was very marked (right-hand graph). Here *t* was found to be 6·77, 9·79, 6·50 and 6·14 for each pair of scores in order from 8 to 60 min. On a two-tailed test in every case $P < 0{\cdot}001$ with 9 *d.f.* (Data for one rat in the second part of the experiment were lost as an airlock had developed in one burette and this was not recognized until after the rat had started to drink.)

If NaCl and water-deprivation exert their effect via the same water-intake controller then the findings of equal drinking in response to a hypertonic injection by both hungry and non-hungry rats rules out the possibility that any continuing neural or hormonal influence was acting multiplicatively on the thirst signal. It does not however rule out the possibility that an inhibitory influence acting subtractively on the water intake controller might be responsible for the reduction of drinking observed where thirst was induced by deprivation of water in the second part of the experiment. It could be argued for instance that in the first part of the experiment, during the 24 hr. deprivation of food, drinking was progressively inhibited so that at the end of this period the rats were in a state of water deficit which was precisely counterbalanced by a subtractive inhibitory signal from the hunger system. Response to the salt injection would on this explanation be the same for hungry and non-hungry rats.

Experiment II was designed to investigate this possibility, and to explore further another unexplained phenomenon observed during Experiment I. This was the disparity between the effects of combined food and water deprivation seen in a 1-hr. drinking test, where drinking was reduced to about a third of the intake following deprivation of water only (see Fig. 1, right-hand graph) and the amount of reduction to only three-quarters seen in the 24-hr. water intake of rats during food deprivation. In 24 hr. the 6 food deprived rats in the first part of Experiment I drank a mean 18·7 ml. as compared with 24·7 ml. for the undeprived animals. The various estimates of reduced drinking made in previous published reports reviewed by Bolles (1961) show a similar puzzling variation, ranging from 29 to 76 per cent reduction of intake during 24 hr. of food deprivation.

In Experiment II a different strain of rats was used to see if this influenced the result, and observations of urine output as well as more accurate measurements of food and water intake on a larger number of rats were made. Furthermore two extra groups were included in the design to investigate the possiblity of a subtractive inhibition. These groups were stomach loaded with water 3 hr. before the salt injection. If hungry rats really were in water deficit which could not be neutralized by drinking because of a subtractive inhibitory signal then one would expect that in the hungry rats this water load would not be released in the diuretic response that noramally follows such hydration, and furthermore the inhibitory signal now being no longer balanced by a deficit would be able to act on the thirst produced by the salt injection. If hungry and non-hungry rats drank the same even after this water load then it could be safely concluded that there is no direct inhibition of thirst by hunger. If some intermediate result were found then it would be possible to calculate the contributions of the inhibition and the reduction of water requirements.

Experiment II

Procedure

The subjects were 32 male albino rats. They were 3 to 4 months old when used and weighed 242 to 330 g. 24 hr. before their drinking test.

The apparatus used was the drinking rack described in Experiment I, modified for use as a set of metabolism cages. Funnels and graduated cylinders collected urine, and a gauze screen in each funnel prevented contamination of the urine with faeces. About 1 ml. of medicinal paraffin was put in each cylinder before urine collection to prevent evaporation. The food provided was always powdered, and the rats reached it by putting his head through a 2½ in. diam. hole in the end of the cage opposite the drinking spout. The top of the food container was level with or slightly below the bottom of the aperture through which the rat put his head, and was fixed ½ to ¾ in. away from it, so that no food could be scratched into the cage. This method was effective since none appeared in the urine cylinders. Each food container was housed in a small cube shaped wire mesh extension bolted on to the cage. It was accessible at the end furthest from the rat via a small sliding aluminium door. Surrounding this food box on all sides except above was a plastic container which collected any spillage. This arrangement ensured that all food not eaten by the animal was recovered with wastage that did not exceed 0·1 g. Food was dried to constant weight at room temperature before use. The burettes used were more accurate than those used in the previous experiment, being graduated to 0·1 ml. The drinking spouts protruded into each cage about ½ in. Below each spout was a small piece of brass tube arranged to catch any drips from the burette and conduct spillage to a cylinder where spillage could be collected under paraffin. This precaution turned out to be largely unnecessary because except on one occasion during a habituation period when a small amount of water was found to have leaked due to only partial closure of a burette tap, no spillage was recorded. Distilled water was used throughout.

The design of this experiment was a 2 × 2 analysis of variance, with independent measurements on 32 rats. The factors investigated in the design were food deprivation (zero or 24 hr.) and stomach load (sham load or 10 ml. distilled water.)

The procedure employed the subjects in batches of 8 at a time (two from each cell of the analysis). Rats were first habituated to the cages for 5 days then either food deprived or not and then tested on the next day. Rats were randomly assigned to a group. A particular group in the design did not always occupy the same cages in different batches. Body weight, food and water intake were recorded every day, and during the second half of the habituation period urine was also recorded. On each of the first 2 days of habituation all subjects were accustomed to receiving salt injections, by receiving intraperitoneally 2 ml. of 1 M NaCl. After an interval of 30 min., during which the rats were placed in a rack of waiting cages without food or water, they were placed back in the metabolism cages and allowed to drink for 1 hr. Amount drunk and latency to drink were noted, and all rats performed satisfactorily, drinking quite promptly by the second occasion and consuming between 2·6 and 7·4 ml. During the second half of the habituation period daily weighing of the animals, etc., was performed at 10-min. intervals, starting at the same time of day.

On the day before testing half the rats were deprived of food. On the test day, exactly 24 hr. since the previous weighing each rat was anaesthetized very briefly in an ether jar, a soft p.v.c. stomach tube with a rounded end was passed. Half the animals were stomach loaded with 10 ml. of water warmed to body temperature, the rest received no water by the stomach tube. Each rat was then put back in his cage with water but no food available. Water intake and urine were recorded every hour for 3 hr., and then the rats were injected at 10-min. intervals with 2 ml. of 1 M NaCl, placed in a waiting cages without food or water for 30 min. and then replaced in the metabolism cages still with no food, to drink for 2 hr. Readings of water consumed were taken at 10-min. intervals for 30 min. and at 1 and 2 hr.

The room temperature range during the experiment was 65° to 78°F. and there was a 12 hr. light–12 hr. dark illumination cycle.

Results and discussion

The main results of this experiment are displayed in Table I. Analysis of variance showed there was no significant difference between the groups in 1 hr. drinking to the second pre-test injection during the habitation period (P for both main groups and the interaction term $>0{\cdot}1$) whereas in their response to the test injection the hungry animals drank slightly but significantly more ($F = 8{\cdot}9$, $d.f. = 1/28$, $P<0{\cdot}01$) than the animals that were not hungry. Stomach loading 3 hr. before the test made no significant difference to the drinking scores ($F = 0{\cdot}34$). The interaction term was also not significant. Far from being inhibited hungry rats thus seem to have their thirst potentiated. Drinking scores after

10 min., 20 min., 30 min. and 2 hr. have also been subjected to analysis and in every case hungry rats drank more than non-hungry ones, all results being significant at the 1 per cent level.

This effect is obviously quite small, since no significant differences between hungry and satiated rats were found in the first experiment in which 12 animals took part. In the second experiment with 32 animals differences significant at the 1 per cent level were found, though these may well have been smaller had not the random assignment of animals to groups (no attempt having been made to match the animals) resulted in slightly but not significantly more drinking in response to the second pre-test injection by rats that were deprived than by those that were not. In any case the increase in drinking was small, and the important observation was that the hungry animals did not drink less.

TABLE I

Scores from each group in Experiment II

	Groups ($N = 8$)	ND NL	ND L	D NL	DL
Measurements taken 24 hr. before stomach tubing	Body wt.	273 (26)	272 (34)	276 (20)	264 (19)
	Previous 24 hr. food	24·0 (2·5)	25·2 (2·7)	25·7 (4·1)	22·1 (1·5)
	Previous 24 hr. water	33·1 (4·3)	35·0 (4·4)	37·2 (6·8)	33·9 (4·9)
	Previous 24 hr. urine	12·8 (2·5)	14·9 (3·7)	14·9 (4·5)	12·7 (5·4)
Measurements taken immediately before stomach tubing	Body wt.	279 (27)	276 (33)	248 (19)	239 (18)
	Previous 24 hr. food	25·4 (2·2)	24·1 (2·3)	0	0
	Previous 24 hr. water	34·5 (3·0)	34·7 (5·6)	28·6 (12·5)	32·8 (18·4)
	Previous 24 hr. urine	11·9 (2·8)	14·0 (4·1)	25·7 (8·1)	28·9 (13·8)
Urine in 3 hr. following stomach tubing		1·4 (0·7)	8·7 (0·9)	1·2 (1·1)	8·7 (0·7)
Drinking in 1 hr. following NaCl injection		5·6 (0·6)	5·3 (2·1)	7·1 (0·8)	6·8 (1·2)

All scores in g. or ml.; groups ND NL = Not deprived—not stomach loaded, ND L = Not Deprived—stomach loaded, D NL = Deprived—not stomach loaded, D L = Deprived—stomach loaded. Numbers in parentheses are standard deviations.

The diuretic response to the water load given by stomach tube on the test day was exactly the same for hungry and non-hungry rats (mean 8·7 ml. for both groups) as compared with 1·4 ml. from the undeprived and 1·9 ml. from the deprived groups that received no stomach load. Absence of any retention of the load indicates that the hungry rats were not in a state of water deficit.

The findings on the drinking tests and urinary responses to the imposed water load make it most likely that hunger directly inhibits thirst in any way. One alternative explanation for the reduced drinking in response to water deprivation by hungry as compared to non-hungry rats, when the hungry rats drink slightly more in response to a salt injection, might be that salt and deprivation act entirely differently. This seems unlikely though Vance (1967) has suggested that "sodium thirst and thirst resulting from normal dehydration are mediated by separate mechanisms" on the basis of some preliminary observations that rats with section of the olfactory bulbs drink less water with a longer latency in response to salt injection than do controls, whereas their *ad lib* drinking is normal.

An explanation along the lines that food deprived rats have less water requirements than when fed, but that despite this, water intake becomes rather labile and somewhat in excess of needs, does deal with the data on drinking in relation to food deprivation. It copes, for instance, with the discrepancy found in Experiment I, between the amount of water drunk during 24 hr, of food deprivation (about three-quarters of normal) and the amount drunk in 1 hr. at the end of 24 hr. of combined food and water deprivation (about one-third of normal). Drinking during the food deprivation period of Experiment II can be seen in Table I, where the mean amount drunk by food deprived animals was only about 85 per cent of control scores.

Two remarkable things occur during food deprivation. One is that the deprived animals excrete roughly twice as much urine as normal, the other is that the variability of the 24 hr. drinking scores increases so that the standard deviation rises, as can be seen in Table I. The rather surprising increase in urine volume in food deprived rats seems first to have been observed by Cumming and Morrison (1960). Though the effect may not be invariable, since there is at least one detailed report (Dicker, 1949) of the precise opposite; a decline in urine volume and water intake. Morrison and his co-workers (Morrison, Mackay, Hurlbrink, Wier, Nick and Millar, 1967; Morrison, 1967, 1968) have recently followed up earlier observations and though they have excluded certain possibilities have not identified with any degree of certainty the cause of the phenomenon. Although one-third of the polyuria can be attributed to diminishing gut contents during starvation the other two-thirds appears to be due to a relative polydipsia. Food deprived rats although they may on average drink less than normal, seem to drink a good deal more than they need.

An explanation that drinking is secondary to a polyuria caused by hunger inhibiting the release of anti-diuretic hormone (ADH), seems to have been ruled out by Morrison *et al.* (1967). They found that injection of ADH into rats during food deprivation restored urine concentration and volume without significantly reducing drinking. The relative polydipsia seems to be primary, and the polyuria is a response to the added water load. Morrison (1968) has also shown that hungry rats will not drink as much if they have to obtain water by CRF bar pressing for very small water reinforcements, or on fixed ratios for reinforcements of 0·1 ml. Also in these cases of high work requirement urine output falls to normal values.

Food deprivation, since it prevents intake of sodium and potassium, might be accompanied by some electrolyte imbalance. If so, then either a relative sodium depletion or a relative potassium depletion are known to cause thirst, and might partly account for the increased drinking. Morrison (1968) puts forward instead a complex hypothesis in which he argues that the apparently excessive drinking maintains an aspect of energy balance. However, neither of these explanations seems satisfactory. The greatly increased variance of *ad lib* water intake scores (see Table I) during food deprivation also argues against a regulatory basis for this kind of drinking, and a non-regulatory basis should be sought. It might be either that drinks are being taken in anticipation of meals that never arrive (drinking just before meals is a well known feature of rats' *ad lib* food and water intake), or that the rats are making an attempt to produce a partial satiety for food by filling the stomach with water. Perhaps the behaviour represents displacement activity, or in some unspecific way the rats are maintaining their normal pattern of water intake even when it is strictly unnecessary.

The whole phenomenon of relative polydipsia during food deprivation recalls the experiments by Falk (1961, 1966, 1967), and several others on schedule induced polydipsia for which food deprivation is a prerequisite. Feeding hungry rats with small quantities of certain types of food (which need not necessarily be dry) on either response contingent or free delivery schedules in which the rate of delivery is less than about one reinforcement per 40 sec. induces polydipsic behaviour. Though the explanation for this is by no means clear it seems not to be a superstitious response since Burks, Hitzing and Schaeffer (1967) have shown that most of the drinking occurs early in each 40 sec. interval on a free delivery fixed 40 sec. delivery schedule; not immediately before the sucrose pellet appears as would be expected if the rats were responding superstitiously. Both schedule induced polydipsia and that which accompanies simple food deprivation seem to be examples of what Falk (1961) recognized as non-homeostatic drinking.

The fact that Morrison's experiments showed that food deprived rats were less inclined to work hard for water recalls another example of apparently non-homeostatic intake, the hyperphagic syndrome (Miller, Bailey and Stevenson, 1950). Animals with ventromedial hypothalamic lesions, though prepared to eat steadily if no effort is required, will not press on high ratios for food, nor will they accept it if it is slightly adulterated. The finding of Mrosovsky (1964) that a similar state of affairs occurs in hibernators, in a way that may be related to food gathering in preparation for hibernation, gives credence to the belief that the effect may reflect some aspect of a normal hypothalamic mechanism and is not merely a bizarre behavioural syndrome resulting from brain damage. If this is true then one is forced to accept that there can be uncoupling of consummatory and appetitive behaviour on some occasions. In the relative polydipsia of hunger the energy devoted to appetitive behaviour in Morrison's (1968) experiment may be more closely related to water need than the amount of free water intake.

These experiments indicate that despite the results of central stimulation studies, hunger does not seem to exert a direct inhibitory influence on thirst. Rather it has somewhat the opposite effect, making the amount of drinking depart from the usual minimal quantity required to replenish water deficits, and producing intake in excess of needs, as well as slightly potentiating the response to a salt injection. The phenomenon of apparently excessive non-homeostatic drinking, as well as the divergences between appetitive and consummatory aspects of hunger and thirst, may well be important in understainding the control of motivated behaviour.

We would like to thank the Medical Research Council for supporting this research. We are also grateful to Mrs. S. A. Larcombe for assistance in the first experiment, which was carried out at the National Physical Laboratory.

References

Adolph, E. F. (1947). Urges to eat and drink in rats. *Am. J. Physiol.* **151**, 110–25.

Bolles, R. C. (1961). The interaction of hunger and thirst in the rat. *J. comp. physiol. Psychol.* **54**, 58–04.

Burks, C. D., Hitzing, E. W. and Schaeffer, R. W. (1967). Drinking response distributions associated with a 4 per cent sucrose F.F.I. schedule. *Psychon. Sci.* **8**, 13–4.

Cizek, L. J. (1959). Long term observations on the relationship between food and water intake in the dog. *Am. J. Physiol.* **102**, 347–52.

Cumming, M. C. and Morrison, S. D. (1960). The total metabolism of rats during fasting and re-feeding. *J. Physiol., Lond.* **154**, 219–43.

Dicker, S. E. (1949). Changes in the extracellular and intracellular fluid phases of muscle during starvation and dehydration in adult rats. *J. Biochem., Tokyo* **44**, 274–81.

Falk, J. L. (1961). The behavioural regulation of water electrolyte balance. *Nebraska Symposium on Motivation*, **9**, 1–33. University of Nebraska Press.

Falk, J. L. (1966). The motivational properties of schedule induced polydipsia. *J. exp. Analysis Behav.* **9**, 19 25.

Falk, J. L. (1967). Control of schedule induced polydipsia: type size and spacing of meals. *J. exp. Analysis Behav.* **10**, 199 206.

Grossman, S. P. (1962). Direct adrenergic and cholinergic stimulation of hypothalmic mechanisms. *Am. J. Physiol.* **202**, 872–82.

Grossman, S. P. (1967). *A Textbook of Physiological Psychology*. New York: Wiley.

Jacobs, H. L. (1964). The interaction of hunger and thirst: experimental separation of osmotic and oral-gastric factors in the regulation of caloric intake. In Wayner, M. J. (Ed.). *Thirst, Proceedings of the 1st International Symposium on Thirst in the Regulation of Body Water*. Oxford: Pergamon Press.

Lepkovsky, S., Lyman, R., Fleming, D., Nagumo, M. and Dimick, M. M. (1957) Gastrointestinal regulation of water and its effect on food intake and rate of digestion. *Am. J. Physiol.* **188**, 327–31.

McFarland, D. J. (1964). Interaction of hunger and thirst in the barbary dove. *J. comp. physiol. Psychol.* **58**, 174–79.

McFarland, D. J. (1965). The effect of hunger on thirst motivated behaviour in the barbary dove. *Anim. Behav.* **13**, 286–92.

Miller, N. E. (1960). Motivational effects of brain stimulation and drugs. *Fedn. Proc. Fedn. Am. Socs. exp. Biol.* **19**, 846–54.

Miller, N. E., Bailey, C. J. and Stevenson, J. A. F. (1950). Decreased "hunger" but increased food intake resulting from hypothalamic lesions. *Science, N.Y.* **112**, 256–9.

Morrison, S. D. (1967). The adrenal cortex and the regulation of water exchange during food deprivation. *Endocrinology*, **80**, 83–59.

Morrison, S. D. (1968). Regulation of water intake by rats deprived of food. *Physiol. Behav.* **3**, 75–81.

Morrison, S. D., Mackay, C., Hurlbrink, E., Wier, J. K., Nick, M. S. and Millar, F. K. (1967). The water exchange and polyuria of rats deprived of food. *Q. Jl exp. Physiol.* **52**, 51–67.

MROSOVSKY, N. (1964). The performance of dormice and other hibernators on tests of hunger motivation. *Anim. Behav.* **12,** 454–69.

OATLEY, K. (1965). *Thirst and drinking mechanisms in the regulation of water intake in rats.* Ph.D. thesis. University of London.

OATLEY, K. (1967). A control model of the physiological basis of thirst. *Med. biol. Engng.* **5,** 225–37.

RADFORD, E. P. (1959). Factors modifying water metabolism in rats fed dry diets. *Am. J. Physiol.* **196,** 1098–108.

VANCE, W. B. (1967). Olfactory bulb resection and water intake on the white rat. *Psychon. Sci.* **8,** 131.

VERPLANCK, W. S. and HAYES, J. R. (1953). Eating and drinking as a function of maintenance schedule. *J. comp. physiol. Psychol.* **46,** 327–33.

Received 31 *January* 1969

Q. Jl exp. Psychol. (1969) **21,** 172–179

JUDGING THE ORDER OF VISUAL STIMULI

KEITH OATLEY, ANTHONY ROBERTSON† AND P. M. SCANLAN
Laboratory of Experimental Psychology, University of Sussex, Brighton and the National Physical Laboratory, Teddington

Psychometric functions obtained for judgements of the order of two long duration light flashes were about half as steep as those previously obtained for short flashes. The range from 50 to 75 per cent correct judgements corresponded to interstimulus intervals from 36 to 98 msec. in the 6 subjects tested. Stimuli ½ degree above and below the fixation point produced slightly more accurate judgements than pairs of stimuli 2 degrees above and below this point. The shape of the curves did not conform to a normal ogive of the kind normally expected in threshold determinations. These data may be explicable in terms of a visual sampling mechanism capable of being triggered by incoming stimuli.

Introduction

Hirsh and Sherrick (1961) in discussing their experiments on perceived order of brief stimuli presented either within or between different sense modalities have stated a general rule for this kind of judgement. They say that for 75 per cent correct order judgements the interval between two stimuli is about 20 msec. and independent of the sense modality used, provided only that the stimuli can be separately identified and named by the subject.

Almost all subsequent reports of temporal discrimination experiments have however either explicitly or implicitly questioned the validity of this generalization. Within visual (Rutschmann, 1966) and tactile (Halliday and Mingay, 1964) modalities position of stimulation can affect apparent simultaneity in a direction, but not necessarily to the full extent, that might be predicted from afferent conduction times. In the auditory modality Babkoff and Sutton (1963) have shown that loudness affects the task. For intermodality judgements involving auditory and visual stimuli apparent simultaneity can occur when a click precedes a flash by some 40 msec. (Rutschmann and Link, 1964). Furthermore, the slope of the psychometric function may be such (Kristofferson, 1967*b*) that the interval between 50 and 75 per cent correct judgements is 2 to 3 times the 20 msec. found by Hirsh and Sherrick.

Although all these findings represent departures from the initial generalization about this type of task perhaps the most surprising new result is that of Robinson (1967). He found that the order of visual stimuli presented dichoptically could be judged correctly on 100 per cent of presentations when the interval between them was only 5 msec. There must be some doubt about Robinson's finding however since Thor (1967) in attempting to replicate this experiment in almost identical conditions found the psychometric function for dichoptic presentation to be very much less steep, with none of the six subjects reaching 100 per cent correct performance at 20 msec. and only one by the 40 msec. interstimulus interval.

---

† Present address: Committee on Mathematical Biology, University of Chicago, Chicago, Ill., U.S.A.

As regards the shape of the psychometric function, most workers have implicitly supported classical threshold theories by plotting their results as straight lines on probability paper. Only Kristofferson (e.g. 1966, 1967*a*, *b*) has thought otherwise, and he regards data from this kind of judgement as providing evidence for a perceptual moment theory of the kind originated by Stroud (1949) and reviewed and given support by White (1963), and Shallice (1964, 1967) among others. In its simplest form this hypothesis states that psychological time is quantized in approximately equal periods (moments) during which only the average of any change of stimulus parameter such as position is known. It is rather as if a sequence of long exposure photographs is taken of the stimulus array. Only the averaged information from each exposure is passed on for further analysis and new data are only available at the end of each exposure. For a correct order judgement to be made stimulus events must therefore fall in separate moments. If no correlation exists between the timing of stimuli and moments, it would follow that the psychometric function relating the frequency of correct order judgement to inter-stimulus interval is a straight line reaching 100 per cent correct responses at an interval equal to the perceptual moment. Kristofferson discusses his version of this hypothesis in terms of being able to switch attention to a new channel only at the end of each moment. He reports that his data give support to this hypothesis, fitting either a straight line of the kind described, or a curve made up of two linear segments, which he explains in an extension of the moment hypothesis by allowing the possibility of the second stimulus falling not into the next moment after that occupied by the first stimulus, but into the one after next.

We have been concerned in the first place to test the validity of Hirsh and Sherrick's generalization for order judgements with visual stimuli for which only the onset is seen by the subject before he makes his judgement. In this way apparent movement is eliminated and furthermore only one event associated with each stimulus can be used in the discrimination. Second we have examined in some detail the shape of the psychometric function under these conditions to see whether any decision can be made between the continuous intake of stimulus information implied by an ogival curve and discontinuous intake such as that supported by the perceptual moment hypothesis and reflected in a linear function.

Method

The subjects, 6 members of N.P.L. staff, were required to view binocularly the screen of an oscilloscope (Tektronix 565 with P2 phosphor) with the graticule illumination turned off, at a distance of 57 cm. A polarized hood was used to prevent reflections from the screen. The oscilloscope was used to display a small (1 mm. diameter) transient fixation light, followed by two long duration stimulus light—spots which appeared, separated by some small interval, one above and one below the fixation point.

Each trial began with the brief appearance of the fixation light in the centre of the screen. Subjects were instructed to fixate this position. 300 msec. after the onset of this warning flash, the first stimulus spot appeared, randomly above or below the fixation point. After a variable interval of between 1 and 100 msec. the second spot appeared, equidistant from, but on the opposite side, of the fixation point. Each stimulus spot was of 2 sec. duaration. Subjects were required to report, before the disappearance of either, which of the stimulus lights appeared first. They responded by saying either "Top" or "Bottom." Subjects had been informed (quite correctly) that the spots never appeared simultaneously, and they were not allowed a "Don't Know" response.

Two conditions were used in the experiment. In one the stimuli appeared 0·5 cm. (0·5 degrees) above and below the position of the fixation point (inter-stimulus separation of 1 cm.), and in the other they appeared 2 cm. above and below the fixation point (inter-stimulus separation of 4 cm.). Both the stimuli and the fixation light were approximately 1 mm. in diameter (6 min. of arc).

Luminance was measured with an SEI photometer. This instrument could not be used accurately for the rather small stimuli in the conditions of viewing (with a polarized hood fitted). The spot brightness was however well above threshold and to an order of magnitude 50 ft. Lamberts. The fixation light was somewhat dimmer than this. Background luminance of the oscilloscope screen was 0·05 ft. Lamberts. Ambient illumination in the experimental room was 0·03 candles/sq. ft.

Each subject completed trials with nine inter-stimulus intervals: 1, 10, 20, 32, 40, 50, 63, 80 and 100 msec. between stimulus onsets. Twenty-five consecutive trials with the same interstimulus interval made up a single block. The first spot appeared at the top randomly on exactly half the total trials presented to each subject. The order of stimulus appearance was set up manually, according to a list derived from random number tables, by the experimenter who sat behind a screen, operating a noiseless switch and recording responses. A single session, lasting about 20 min., involved the subject in 9 blocks of 25 trials, one block at each inter-stimulus interval. The order of blocks in a session was randomized. Within a single session the stimulus separation was fixed at either 1 or 4 cm., and the order of these conditions was randomized. Subjects were given no immediate knowledge of results, but were allowed to ask about their performance at the end of each block if they wished.

Stimuli were derived from Tektronix waveform and pulse generators. The warning and fixation spot was produced by transient Z-modulation of one of the two C.R.T. electron guns. Each stimulus spot above or below the fixation spot was produced by a 2 sec. pulse (50 V. gate output with measured rise time 2 μsec.) starting below the oscilloscope screen and having an amplitude to coincide with the desired position on the screen. The top and bottom spots were transmitted through two channels of a four-channel plug-in (type 3A74) and were of equal brightness.

Results

Of the 6 subjects, 3 completed a total of 200 trials at each of the nine intervals; 2 completed 100 trials at at each interval; the sixth completed 200 trials at nine intervals, but early in the experiment it was clear that she was nowhere near the 100 per cent correct level at 100 msec. and was given a further 200 trials each of 130 and 160 msec. intervals. For each subject a three-way analysis of variance, each cell having one observation (the number of correct responses out of 25) has been performed. In each analysis the factors were order, spot separation and spot delay. In two subjects (A.R. and R.G.) there were significant order effects, presumably due to practice at the task, and in another (P.M.) the order times spot separation interaction was significant.

The analysis of variance indicated that the effect of retinal position of stimuli (0·5 or 2 degrees from the centre of the fovea) was significant in only 1 subject

(P.T.) at the 5 per cent level. In 4 other subjects there was a slight effect, ranging in significance as judged by the F ratio from P approx. $= 0{\cdot}25$ to approx. 0·1. There were no interactions with this factor and in these five subjects the effect was in the same direction with the curve for the spots separated by 1 cm. running parallel to and above those for 4 cm. separation. This must mean that there is some slight increase in the variability of conduction time from extra-foveal as compared with near-foveal regions of the retina.

The effect of inter-spot interval was as expected highly significant for all subjects. Except in the cases described above no effects were significant at the 5 per cent level and so we have lumped the results of all sessions together and these are displayed as a function of inter-stimulus interval for each subject in Figure 1.

In each of the graphs confidence limits for the binomial distribution are shown for $P = 0{\cdot}05$.

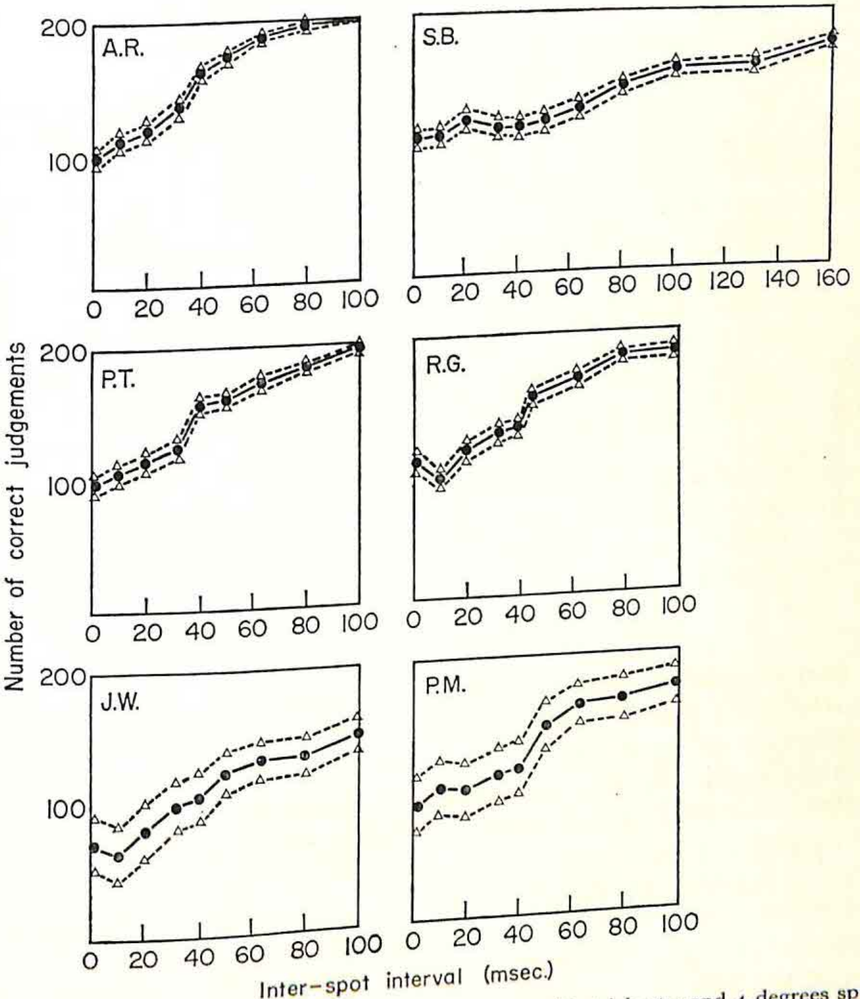


FIGURE 1. Individual results (solid lines) for all subjects, with trials at 1 and 4 degrees spot separation lumped together. Triangles joined by dotted lines mark fiducial limits for the binomial distribution at $P = 0{\cdot}05$.

Discussion

Our results, like those of most other experimenters on the perception of serial order, have failed to support the generalization of Hirsh and Sherrick (1961) that about 20 msec. are invariably needed for the 75 per cent correct discrimination of temporal order. For the six subjects in this experiment the interval between 50 per cent and 75 per cent correct judgement ranged from 36 to 98 msec. There are two possible reasons for this departure. Only one event, the stimulus onset, was available for judgement, and also no apparent movement effects were produced by these stimuli. Though the apparatus that we used was not really suitable for giving short stimuli since the offset of the stimulus was relatively slow and dependent on the decay of the oscilloscope phosphor, we did test one subject (A.R.) with spots of approximately 10 msec. duration and found both that apparent movement was the one obvious feature of the display, and that the slope of the psychometric function was about twice that for the longer stimuli and within the range obtained by others in similar tasks. Apparent movement as such may be responsible for this difference implying a different mechanism (cf. Barlow and Levick, 1965) for temporal discrimination in this situation. However, the slopes found are in the same range as those described by Kristofferson (1967) who is the only other investigator we know to have deliberately allowed only one event (in his case the stimulus offset) to be used in the judgement, although his experiments were with light-sound pairs.

The results of Robinson (1967) would, if substantiated, seem to require a separate category of temporal judgement, in which for dichoptic viewing very short intervals of less than 5 msec. can always be accurately discriminated. Even before Robinson's report a similar ability was well known in the auditory modality where intervals of much less than 1 msec. can be resolved in sound localization (Rose, Gross, Geisler and Hind, 1966). Bekesy (1967) has also described judgements of localization to the site of the first of two successive stimuli in gustatory, olfactory and tactile modalities. In all cases successful discriminations were made with inter-stimulus intervals of 2 msec. or less, and in the tactile modality this ability was not dependent upon stimulation being presented separately to two sides of the body. Since Thor (1967) has failed to confirm Robinson's result there must remain some doubt as to whether these very short interval discriminations can be made in the visual system.

The question of the slope of the psychometric functions obtained in experiments on temporal discriminations will probably need careful parametric examination before the precise significance of differing results from different laboratories can be determined. At least as important, and almost undiscussed is the shape of the psychometric functions obtained. Though Kristofferson (1966, 1967*b*) fits his data to straight lines he does not allude to any of the previous work in this area, and no other workers make comment on the shape of the psychometric functions they obtain. On the other hand except for Thor (1967) all others who present their results graphically make the assumption that the best fit is a normal ogive, since they present their results on Gaussian probability paper. This curve plotted on linear co-ordinates is as shown in Figure 2(a). For many purposes only the upper half of the graph, between chance (50 per cent correct on a binary decision) and errorless (100 per cent correct) performance is needed.

This ogival curve derived from classical psychophysical theory, is regarded as being similar to that obtained from any type of difference threshold. A fundamental property is that the slope relating correct peformance to the stimulus is at its steepest where the two stimuli to be differentiated are in fact equal.

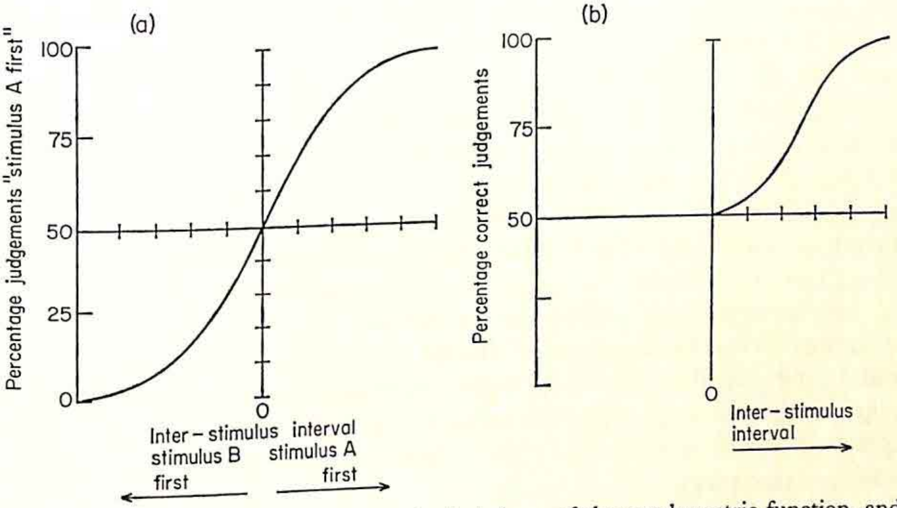


FIGURE 2. The curve (a) represents the classical shape of the psychometric function, and that assumed by authors of most previous experiments. The curve is a normal ogive with its centre intersecting with the 50 per cent correct level. For two stimuli A and B the responses of the subject are plotted as percentage of the subject's reports of "stimulus A first." The curve (b) is an idealization of the results obtained in this experiment. The curve is sigmoidal; it may also be an ogive, but it starts at the 50 per cent level rather than having its midpoint there. Clearly it can also extend into the lower left segment if responses are plotted as in (a) but in this case the whole curve would be composed of two ogives and not one.

This is not by any means obvious intuitively. In many types of difference threshold one would have imagined that there was a lower limit: the "just noticeable difference." Furthermore it may not be the case empirically. In some of the published work on temporal discrimination where enough observations have been made and the inter-stimulus intervals have been sufficiently closely spaced the shape of the curve can be seen to depart from the normal ogive.

We have for instance tried fitting the individual results of this experiment and those of Kristofferson (1967*b*) to straight lines on probability paper, but in only a few cases are the fits acceptable. The curve suggested by the majority of our results is sigmoid or S-shaped. It has the general form of Figure 2(b), with ends which, as they approach the 50 and 100 per cent levels, become less steep than the middle section.

The function determined by Thor (1967) for dichoptic signals is also like Figure 2(b), as are most of Kristofferson's (1966) results. Halliday and Mingay's (1964) histograms may be similar, although in their case the results are more difficult to interpret since they allowed their subjects a third category of "stimuli judged simultaneous." Other reports of temporal discrimination experiments are insufficiently detailed to allow an accurate assessment of the shape of the curve, though Rutschmann's (1966) results look as if they might be best fitted by a normal ogive, in the classical way.

If the psychometric function in this kind of experiment is S-shaped as in Figure 2(b), then there seem to be two possible explanations. If the intake of stimulus information is continuous, then there may be a threshold interval so that when two stimuli are presented within this they are nearly always perceived as simultaneous. In the case of our subjects this threshold interval would be of the order of 40 msec. The explanation seems unattractive because it would require a very large amount of unexplained variability in the process. From our results the standard deviation of this variaiblity would be 20 msec. or more, and has to be compared with the very remarkable accuracy that is displayed in auditory localization phenomena (e.g. Rose *et al.*, 1967).

A second type of theoretical basis would be that of perceptual moments. Kristofferson (1966, 1967*a*, *b*) favours this, but it does not seem obvious that his individual results fit either the one or the two linear segments that his hypothesis allows. In particular his (1967*b*) data miss out vital observations in the region of 50 per cent correct responses. Without this information the proposition that the underlying distribution is a straight line rising from 50 per cent correct to about 85 per cent correct and then inflecting into a less steep straight line rising to intersect with the 100 per cent correct level must remain insubstantial.

If a sampling mechanism is responsible for the shape of the psychometric function, the results shown in Figure 1 suggest that moments are triggered by the stimulus. Such a hypothesis has been suggested elsewhere (e.g. Venables, 1960; Latour, 1967). Their formulation needs modification before it will fit these data, however. We would propose that a stimulus can trigger a moment, which would close, not immediately, but after a delay of 40 to 50 msec. As in other moment theories order judgements can only be easily made if stimuli occur in different moments. A possible exception is that judgements might sometimes be made on the basis of different brightness of the two stimuli averaged over a moment. Without a triggering stimulus the moment timing mechanism would continue to take in samples at its own "freerunning" rate. The initial shallow slope of the curve can then be explained by the first stimulus arriving within 50 msec. of a moment closure not caused by external triggering, so that a second stimulus does not have to wait as long as 50 msec. to get into the next sample. The steep part of the curve occurs where the first stimulus successfully triggers the sample closure so that if the second stimulus comes more than 50 msec. later it is caught in the next moment. The decline in steepness of the curve thereafter may indicate some sort of uncertainty in the mechanism, if for instance the first stimulus is not always sufficient to trigger closure of a moment. This delay system would allow the discrimination mechanism time to store or dispose of any computation going on at the arrival time of a stimulus. Although it has a much shorter time scale this hypothesis has similarities to the psychological refractory period hypothesis advanced by Welford (1952) and others to explain serial reaction times.

Such a model obviously requires further experimental test; and some predictions that it makes for simple reactions to pairs of successive stimuli are being explored.

Whether the intake of stimulus information is continuous or intermittent, it does seem to be the case that the type of psychometric function obtained from temporal discriminations is not always the classical one usually assumed, and that its shape may give valuable information about the temporal characteristics of perceptual processes.

Some of this work was part of the research programme of the National Physical Laboratory. The computing facilities were made available at the Laboratory of Experimental Psychology, University of Sussex by a grant from the SRC.

References

Babkoff, H. and Sutton, S. (1963). Perception of temporal order and loudness judgements for dichotic clicks. *J. acoust. Soc. Am.* **35**, 57–47.

Barlow, H. B. and Levick, W. R. (1965). The mechanism of directionally sensitive-units in rabbit's retina. *J. Physiol., Lond.* **178**, 477–504.

von Bekesy, G. (1967). *Sensory Inhibition.* Princeton, N.J.: Princeton University Press.

Halliday, A. M. and Mingay, R. (1964). On the resolution of small time intervals and the effect of conduction delays on the judgement of simultaneity. *Q. Jl exp. Psychol.* **16**, 35–46.

Hirsh, I. J. and Sherrick, C. E. (1961). Perceived order in different sense modalities. *J. exp. Psychol.* **62**, 423-32.

Kristofferson, A. B. (1966). Further experiments on successiveness discrimination. *N.A.S.A. Report No.* CR4–54.

Kristofferson, A. B. (1967*a*). Attention and psychophysical time. *Acta Psychol.* **27**, 93–100.

Kristofferson, A. B. (1967*b*). Successiveness discrimination as a two-state quantal process. *Science, N.Y.* **158**, 133–79.

Latour, P. L. (1967). Evidence of internal clocks in the human operator. *Acta Psychol.* **27**, 341–8.

Robinson, D. N. (1967). Visual discrimination of temporal order. *Science, N.Y.* **156**, 126–34.

Rose, J. E., Gross, N. B., Geisler, C. D. and Hind, J. E. (1966). Some neural mechanisms in the inferior colliculus of the cat which may be relevant to the localization of a sound source. *J. Neurophysiol.* **29**, 288–314.

Rutschmann, J. and Link, R. (1964). Perception of temporal order of stimuli differing in sense mode and simple reaction time. *Percept. Mot. Skills.* **18**, 345–52.

Rutschmann, R. (1966). Perception of temporal order and relative visual latency. *Science, N.Y.* **152**, 1099–101.

Shallice, T. (1964). The detection of change and the perceptual moment hypothesis. *Br. J. Statist. Psychol.* **17**, 11–35.

Shallice, T. (1967). Temporal summation and absolute brightness thresholds. *Br. J. Math. Statist. Psychol.* **20**, 129–61.

Stroud, J. M. (1949). The psychological moment in perception. In (Ed.) *Trans. 6th Conference on Cybernetics.* Van Foerster. New York: Josiah Macy.

Thor, D. H. (1967). Dichoptic viewing and temporal discrimination: an attempted replication. *Science, N.Y.* **158**, 170–45.

Venables, P. H. (1960). Periodicity in reaction times. *Br. J. Psychol.* **51**, 37–43.

Welford, A. T. (1952). The 'Psychological refractory period' and the timing of high-speed performance—a review and a theory. *Br. J. Psychol.* **43**, 2–19.

White, C. T. (1963). Temporal numerosity and the psychological unit of duration. *Psychol. Monog.* **77**, No. 12, Whole No. 575.

Received 27 *October* 1968

Q. Jl exp. Psychol. (1969) **21**, 180-184

TACTILE SHORT-TERM MEMORY

ELIZABETH Q. GILSON AND A. D. BADDELEY
Laboratory of Experimental Psychology, University of Sussex

Subjects tried to recall the location of a tactile stimulus on the underside of the forearm after delays of 0, 3, 5, 10, 15, 30, 45 and 60 sec. When "rehearsal" was prevented by requiring subjects to count backwards during the delay, accuracy of recall decreased systematically reaching an asymptote after 45 sec. When subjects were left free to "rehearse," this did not affect the decline in accuracy over the first 10 sec. Between 10 and 15 sec. there was a significant *increase* in accuracy followed by a slow decline which had not reached asymptote by 60 sec. It is suggested that tactile STM (short-term memory) depends on two processes, a fading sensory trace which is unaffected by distraction and a less labile system which does not appear to be verbal but which depends on "rehearsal."

Introduction

Melton (1963) has suggested that the traditional distinction between long-term memory (LTM) and short-term memory (STM) may be unnecessary, since a single set of assumptions, those of interference theory, could explain all the known phenomena. This assertion stimulated a good deal of research, much of which is difficult to reconcile with the hypothesis of a unitary system, and suggests that we must distinguish between at least two processes, one showing stability over a minimum of several minutes, and a second, which is characterized by rapid forgetting reaching an asymptote after a matter of seconds (Baddeley, 1966; Glanzer and Cunitz, 1966; Milner, 1966; Peterson, 1966). Virtually all the evidence for a separate STM system comes from the retention of verbal materials, however, a type of memory which appears to depend largely on acoustic coding. The possibility remains therefore that other types of coding may show quite different rates of forgetting indicating a multiple memory system system rather than a simple STM-LTM dichotomy.

One way of investigating this possibility is to study the retention of material which cannot readily be acoustically coded. If the simple dichotomy found in verbal learning is typical, one might expect to find marked forgetting over the first 20–30 sec. If such forgtting does not occur, it would suggest a multiple system with each modality having its own characteristic rate of forgetting.

There have been three recent studies of STM for non-verbal material. The first of these (Adams and Dijkstra, 1966) studied kinaesthetic STM. Subjects attempted to reproduce the movement of a slide along a rod after an unfilled interval of 0 to 120 sec. Adams and Dijkstra found rapid forgetting reaching an asymptote after about 80 sec., and although they noted a number of differences between their results and comparable data from verbal learning, their results did suggest a kinaesthetic STM with roughly the same time constant as verbal STM.

Comparable data were obtained by Posner and Konick (1966) using a lever positioning task, although it is difficult to draw any conclusions about the point of asymptote from Posner and Konick's data since the longest delay they used was 30 sec. This problem occurs also in their study of STM for visual location, for which the longest delay was 20 sec. A more detailed study of STM for visual location has been made by Dale (personal communication), who studied the reproduction of the location of a dot in a square after intervals of 0 to 64 sec. under a range of conditions. The most marked forgetting occurred with a brief (0·5 sec.) exposure of the stimulus and an interval filled with digit cancelling, and reached an asymptote after about 30 sec.

Thus, such data as exist are consistent with the assumption that STM systems showing forgetting over a matter of seconds are not limited to verbal material. The material from other modalities is, however, still very sparse and is certainly not enough to justify any general statements about STM. The present study therefore aims to extend our knowledge of STM to a further modality, that of touch.

Method

Subjects were touched at a point on a line drawn on the underside of their forearms, and subsequently tried to indicate the point stimulated after a delay of 0 to 60 sec. Half the subjects were required to count backwards during the delay, while the rest were free to concentrate on remembering the stimulus.

Procedure

A scale in 1 cm. units was drawn in ink on the underside of the subject's right arm extending from below the elbow to the wrist (about 2 in. above the plam). The arm was laid flat on a table. The subject sat at right angles to the table with his arm outside his field of view. Subjects were also instructed not to look at their arms during the experiment, and most achieved this by closing their eyes.

Since subjects' arms differed in length, the number of 1-cm. divisions varied ranging from 14 to 22. However, to minimize use of the elbow and wrist as anchor points only the eight most central points were used. In each of 10 trial blocks, each of these eight locations was stimulated once. Each trial block comprised one occurence of each of eight delays, 0, 3, 5, 10, 15, 30 45 and 60 sec. The order of delays was randomized in each block. The eighty trials took approximately 50 min. including a 5-min. break in the middle.

The arm was touched lightly using a medium-point ball-point pen, and recall, which was cued verbally, involved pointing to the location of the touch with the finger, since this appears to be most compatible response (Sinclair, 1967, p. 177).

Twenty right-handed undergraduates served as subjects. Half served in condition R (rest) and were asked to sit in silence during the retention interval and to "concentrate on trying to remember" the position touched. The other half performed condition A (activity) and were required to count backwards at their own pace in threes, from a three-digit number spoken by the experimenter as she touched the subject's arm. This is a standard procedure for minimizing rehearsal in verbal learning studies (Peterson and Peterson, 1959). A similar task has been found to influence visual though not kinaesthetic memory (Posner and Konick, 1966).

Results

Recall responses were measured to the nearest 0·5 cm. using the scale drawn on the subject's arm. Mean error in millimetres for the two conditions as a function of delay is shown in Figure 1.

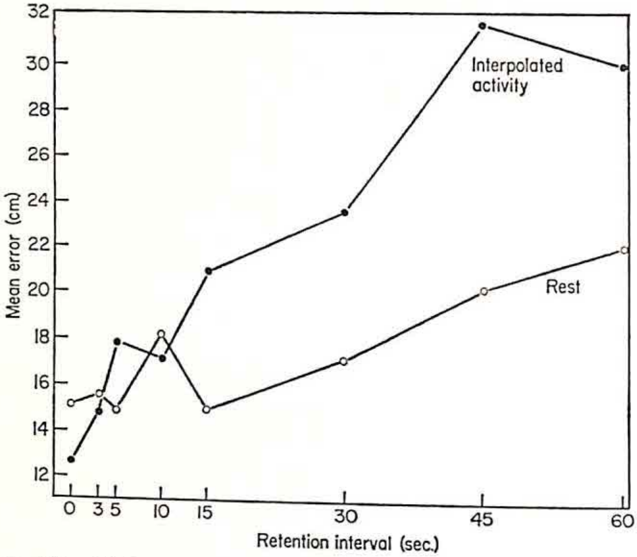


FIG. 1. Accuracy with which a tactile stimulus was recalled as a function of delay and activity during the retention interval. (●) Recall after a distracting task; (○) recall with an unfilled retention interval. Each point based on 100 observations.

Analysis of variance showed significant effects of subjects (F (19, 126) = 14·4, ($P<0·001$), delay (F (7, 126) = 28·3, $P<0·001$) and a significant interaction between delay and activity (F (7, 126) = 8·4, $P<0·001$). The main effect of activity did not, however, reach significance (F (1, 18) = 3·2, $P = 0·10$). The reason for this is clear from Figure 1; comparison using the t-test at each delay showed no difference between the two conditions over the first 10 sec. of retention, but a significant difference at each delay beyond 10 sec. ($P<0·02$ in each case).

In the activity condition when subjects are prevented from "rehearsing," the results are quite straightforward with a systematic increase in error which appears to have reached an asymptote by about 45 sec. When no activity is required during retention the picture is more complex. For the first 10-sec. sec. there is a significant *improvement* in performance which is shown by 9 of the 10 subjects and is statistically significant on a Wilcoxon Test ($T = 3$, $N = 10$, $P = 0·01$, two tail). This significance level should of course be regarded with caution since it is based on a single selected *post hoc* comparison. From this point onwards subjects in condition R show a gradual decrement in performance which shows no sign of levelling off within 60 sec.

Studies of verbal STM have demonstrated that the first item presented shows much less forgetting than subsequent items (Keppel and Underwood, 1962), an effect which is analogous in some respects to proactive inhibition (PI) in LTM. A study in which several delays are used is not ideal for studying PI in STM since trial number and delay are inevitably confounded. It is, however, possible to obtain a rough indication of whether or not the effect occurs if performance on the first item presented is compared with the next item having the same delay. This was done separately for conditions A and R. Condition A, where "rehearsal" was minimized showed no PI effect, five subjects were better on the first presentation and five worse. In condition R, however, performance on the first item was significantly better than on the next equivalent item using the Wilcoxon test ($T = 6$, $N = 9$, $P < 0.05$ two-tail).

Discussion

When subjects are prevented from "rehearsing," retention of the location of a touch stimulus shows a rate forgetting of the same order of magnitude as that found in verbal STM, reaching an asymptote after about 45 sec. Unlike verbal STM, however, the first trial does not seem to produce any less forgetting than subsequent trials, so far as the evidence goes.

When there is no interpolated task, however, the overall rate of forgetting is slower and subjects still appear to be forgetting after 60 sec., the longest delay. A simple assumption of slower forgetting, however, is not enough to account for the data. The absence of a significant effect of activity, the significant activity delay interaction, and the significant improvement in performance between 10 and 15 sec. shown by the "rehearsal" group all suggest a more complex picture, possibly involving two separate processes. The first of these seems to be unaffected by the counting task since performance in the two conditions does not differ. Between 10 and 15 sec. this appears to be replaced by a second process which produces an immediate improvement in performance followed by somewhat more gradual forgetting which has not levelled out at the longest interval of 60 sec. Such a conclusion gains some support from subjects' introspective reports. These suggest that stimulation was followed by a sensory after-image consisting of a tingling sensation, which sometimes became clearer over the first 3 sec., but which then began to fade until it had virtually disappeared after 10 sec. This after-image is presumably responsible for the performance of both groups over the first 10 sec. and is probably the only source of information available to subjects in condition A. The fact that performance does not asymptote until 45 sec. have elapsed suggests that the information has not completely disappeared by 10 sec. as some subjects suggested. The absence of a PI effect in this condition is presumably due to the fact that very little information persists from one trial to the next.

Subjects in condition R, instructed to "concentrate on remembering," report that after having "no idea" of the location of the stimulus after 10 sec., by 15 sec. they "somehow knew" where they had been touched, although none seemed to have any idea how they knew. The cue did not appear to be verbal since the only subject claiming to use verbal labels was one with an atypically short arm containing only 14 possible locations. She claimed to have used a visual image of scale together with category names; her performance level was slightly below the mean. One possible explanation is suggested by Melzack and Eisenberg's (1968) observation that a touch on the lip or cheek may produce a tickling sensation which fluctuates in strength and may last up to half an hour. However, Melzack did not find this effect on the eyelid, the finger or the back of the hand and it thus seems unlikely that it would be found on the forearm. Second, no such effect was reported by our subjects. Finally, the Melzack effect was not influenced by distracting the subject during the retention interval, whereas the effect we observed was clearly dependent on the absence of a distracting task. While the nature of tactile "rehearsal" presents an intriguing problem it would seem wise to delay speculation until the basic phenomena have been more firmly established. It is particularly necessary to replicate the improvement in performance shown between 10- and 15-sec. delays.

In general, these results suggest that when "rehearsal" is minimized, forgetting occurs over a matter of seconds, and appears to reach an asymptote after a delay of roughly the same order of magnitude as is found for verbal, visual and kinaesthetic STM. As in visual and verbal memory, when "rehearsal" is encouraged forgetting is slower, though the nature of tactile "rehearsal" remains obscure.

This report is based on an experiment submitted by the senior author in partial fulfilment of the requirements governing the degree of B.Sc. (Experimental Psychology) at the University of Sussex. The authors are grateful to Miss D. Scott and Mr. D. Cooke for statistical assistance and advice. The second author is grateful to the Medical Research Council for financial support.

References

ADAMS, J. A. and DIJKSTRA, S. (1966). Short-term memory for motor responses. *J. exp. Psychol.* **71**, 314–8.

BADDELEY, A. D. (1966). The influence of acoustic and semantic similarity on long-term memory for word sequences. *Q. Jl exp. Psychol.* **18**, 302–9.

GLANZER, M. and CUNITZ, A. R. (1966). Two storage mechanisms in free recall. *J. verb. Learn verb. Behav.* **5**, 351–60.

KEPPEL, G. and UNDERWOOD, B. J. (1962). Proactive inhibition in short-term retention of single items. *J. verb. Learn. verb. Behav.* **1**, 153–61.

LOESS, H. (1964). Proactive Inhibition in short-term memory. *J. verb. Learn. verb. Behav.* **3**, 362–8.

MELTON, A. W. (1963). Implications of short-term memory for a general theory of memory. *J. verb. Learn. verb. Behav.* **2**, 1–21.

MELZACK, R. and EISENBERG, H. (1968). Skin sensory afterglows. *Science, N.Y.* **159**, 445–7.

MILNER, B. (1966). Amnesia following operation of the temporal lobes. In WHITTY, C. W. M. and ZANGWILL, O. L. (Eds.), *Amnesia*. London; Butterworths.

PETERSON, L. R. (1966). Short-term verbal memory and learning. *Psychol. Rev.* **73**, 193–207.

PETERSON, L. R. and PETERSON, M. J. (1959). Short-term retention of individual verbal items. *J. exp. Psychol.* **58**, 193–8.

POSNER, M. J. and KONICK, A. F. (1966). Short-term retention of visual and kinaesthetic information. *Organiz. Behav. hum. Perf.* **1**, 71–86.

SINCLAIR, D. (1967). *Cutaneous Sensation*. London: Oxford University Press.

Received 3 February 1969

Q. Jl exp. Psychol. (1969) **21**, 185–191

RETRIEVAL TIME AS A FUNCTION OF MEMORY ENSEMBLE SIZE

G. E. BRIGGS AND J. M. SWANSON

Ohio State University

A direct test was performed of Oldfield's (1966) derivation of the relationship between naming response latency and memory ensemble size. The results confirmed Oldfield in every particular. Further, the results support Sperling's (1967) suggestion of a recognition buffer-memory mechanism between the initial stimulus-encoding stage and the long-term memory stage of human information processing.

Introduction

Oldfield (1966) addressed himself to the organization of human memory in the first Sir Frederic Bartlett Lecture. He reasoned that more commonly used words should be more readily accessible from memory than those words employed less frequently, and he suggested that Zipf's Law (Zipf, 1935) provides a key to this organization. The latter states that in a large sample of words there is an inverse linear relationship between the log of the number of words with frequency of occurrence f and $\log f$. Thus, one may conceive of memory for object names as organized into ensembles of names with the largest ensemble for least commonly used names, a somewhat smaller ensemble for more commonly used names, and so on. If memory search involves first the selection of an ensemble and then a dichotomous search of the ensemble, it follows that object naming can proceed efficiently and at a rate well within empirical estimates of skill limits involving linguistic material.

Oldfield determined that a Zipf-type relationship held for object names including those used earlier by Oldfield and Wingfield (1965) in a study of response latency for naming pictured objects as a function of name frequency of occurrence in textual samples. At that point Oldfield had two sets of data: (a) naming response latency as a function of log frequency of word occurrence; and (b) the Zipf relationship between these and other equally likely words and the log frequency of word occurrence. Given two variables expressed as a function of a common third variable, it was a simple matter to determine the relationship between the first two variables. The fitted equation was $RT = 373 + 58 \log_2 M$ where RT is average naming reaction time in msec. and M is estimated memory ensemble size.

The following interpretation of this relationship was offered by Oldfield: (a) the intercept constant, 373 msec., was the average time required to encode the pictorial stimulus information *plus* the time required to decide which memory ensemble to search; (b) the inverse of the slope constant, 17 bits per sec., was the rate of search through the memorial representations within the ensemble selected;

and (c) the latter search was not *seriatim* but presumably involved a sequence of dichotomous eliminations.

The purpose of the present study was to determine more directly the relationship between naming response latency and memory ensemble size. To do so it was obvious that one must control experimentally the frequency of occurrence of stimulus names, and this excluded use of previously familiar stimuli. As such, random figures were selected to be stimulus items (Vanderplas, Sanderson and Vanderplas, 1965), and each was assigned a unique name consisting of a letter (A, E, I or O) and a two-digit number (an even number between 20 and 98).

There were four ensembles (M) developed consisting of 2, 4, 8 and 16 stimulus-response pairs; thus, $\log_2 M = 1, 2, 3$ and 4. A Zipf-type relationship was applied to the exposure of the stimulus-response pairs throughout the experiment such that each item in ensemble $M = 2$ occurred eight times for every occurrence of an item in the $M = 16$ ensemble, four times for every occurrence of an item in the $M = 8$ ensemble, twice for every occurrence of an item in the $M = 4$ ensemble, and so on.

It was predicted that a linear relationship would be found between log ensemble size and vocal reaction time in a paired-associates task which required the subjects to make a naming response. It was predicted further that continued practice at this task would result in a decrease in the slope constant of the above relationship such that this index of memory retrieval time would approach the estimate of 17 bits/sec. derived by Oldfield for pre-experimentally familiar material. Finally, it was predicted that if in a special test the subjects were informed of the ensemble from which a stimulus figure was to be selected just prior to exposure of the figure, the intercept of the reaction time-$\log_2 M$ relationship would be decreased but the slope constant would be unaffected compared to the standard situation wherein the subjects had no advanced information. This latter prediction was based on Oldfield's interpretation, noted above, that the intercept is an index, in part, of the time required to determine which ensemble to search. If this prediction is confirmed, the difference in intercept for the special test data and the normal test data (no advanced ensemble information) would be an indication of the ensemble selection time for these materials.

Method

Materials and procedure

Four ensembles (M) of random figures were established. For half the subjects the $M = 2, 4, 8$ and 16 ensembles consisted of 2 four-sided, 4 six-sided, 8 eight-sided and 16 ten-sided figures, respectively; for the other half of the subjects the $M = 2, 4, 8$ and 16 ensembles consisted of 2 ten-sided, 4 eight-sided, 8 six-sided and 16 four-sided figures, respectively. The random figures were selected from a much larger sample provided by Vanderplas *et al.* (1965) so as to equate association values across ensemble size. All figures used had association values between 20 and 45 per cent.

The names assigned to the two figures in the $M = 2$ ensembles each started with the letter A; for the figures in the $M = 4, 8$ and 16 ensembles each name started with the letters E, I and O, respectively. Thus, the ensembles were distinguished on the stimulus aspect by sidedness and on the response aspect by initial letter in the assigned names. Within ensemble $M = 2$ each of the two figures were assigned an even two-digit number between 20 and 28, such that the full name for a given figure might be A-22. The full name for a figure in the $M = 4$ ensemble consisted of the letter E followed by an even two-digit number between 30 and 38. The digit portions of assigned names for the figures in the $M = 8$ ensemble were selected from even numbers between 40 and 58, while those for the $M = 16$ ensemble were even numbers between 60 and 98. Therefore, each of the 30 figures experienced by a given subject had a unique full name but the initial letter identified the appropriate ensemble and thus was common to all members of that ensemble.

The subjects experienced six daily sessions of preliminary training prior to the start of the data collection sessions. On the initial preliminary session a subject viewed each of the 30 random figures with which he would be working one at a time in a random order. He was instructed simply to respond with the ensemble letter as soon as a figure appeared on the screen. Thus, those subjects who were to experience four-sided figures in ensemble $M = 2$ responded in the initial session with the response "A" whenever either of the 2 four-sided figures appeared; likewise, they responded with "E" whenever a six-sided figure occurred, and so on. The purpose of the first session was to introduce all 30 figures and their assigned ensemble groupings.

In the second preliminary session and continuing through the fifth session, the subjects encountered the 30 figures and their assigned full names in a series of paired-associates lists. Each list consisted of eight stimulus-response pairs, two from each of the four ensembles, and there were eight such lists. Across the eight lists the two $M = 2$ figures appeared eight times for every single occurrence of a figure from the $M = 16$ ensemble, and so on, to establish and maintain the Zipf-type relationship between ensemble size and familiarity. The subjects learned two lists to one perfect trial each in each session. The sixth and final preliminary session involved the same paired-associate procedure, but a subject saw only the $M = 8$ and $M = 16$ ensemble items. This violated the Zipf-type relationship, of course, but it was felt necessary to avoid unusually large numbers of errors in naming responses for these two largest ensembles during the experimental trials to follow. The latter utilized strict adherence to the Zipf relationship, so it was felt that any imbalance from preliminary session six would be corrected quickly in the subsequent part of the experiment.

Following the preliminary training sessions, there were six daily sessions followed by a special test session. In each of the six sessions (to be identified as the differential learning sessions) a subject experienced all 30 stimulus-response pairings in a random sequence of 64 such pairings. Each of the $M = 16$ ensemble items appeared once, each of the $M = 8$ items occurred twice, while each of the $M = 4$ and $M = 2$ items appeared 4 and 8 times, respectively. As in the preliminary sessions the subjects attempted to anticipate the correct name for each random figure in a paired-associates procedure. A different random sequence of the 64 items was employed every 2 days. Within each differential learning session there were two exposures of the 64-item list; the data from the second exposure were utilized in the analysis to be reported.

In the final, special test session each subject first viewed a sample of the figures one at a time and he responded as in the six differential learning sessions by giving the full name for each figure. Following this normal testing procedure, the subjects viewed another sample of figures, one at a time; however, to the second sample the subjects gave only the two-digit number portion of the assigned name. Just prior to each figure presentation the experimenter identified the ensemble from which that figure was to be drawn, e.g. "the next figure will be an 'I' figure."

Throughout the experiment the subjects were instructed to be as fast and as accurate as possible in responding. Further, starting with the fourth session of the differential learning period, a bonus system was used to emphasize the speed and accuracy instructions. The experimenter recorded every response and the vocal reaction times throughout the experiment.

Equipment and subjects

Two slide projectors were employed, one to present the stimulus figures, the other to present the assigned responses on a ground-glass screen. Interval timers were used to control stimulus onset and duration (4 sec.) and response onset and duration (2 sec.). Leaf-type shutters were placed in front of the projectors to provide these exposures. A photosensitive diode was placed in front of the stimulus shutter which activated an electronic counter upon stimulus onset. The counter accumulated time at the rate of 10,000 cts. sec. The subjects responded over an intercom system, which response activated a Schmitt trigger to terminate the electronic counter. The experimenter heard and recorded these vocal responses over the intercom and he could communicate with a subject, when appropriate, over the same system. The subjects heard a constant-intensity white noise over the intercom except when the experimenter was communicating with him. This provided for masking of extraneous laboratory sounds.

Twelve undergraduate students participated in the study, eight males and four females. All responded to an advertisement in the campus paper. None had served in a similar experiment and none had seen the random figures prior to their service. Each received a base pay of $1.25 per session plus a bonus starting with the fourth session of differential learning.

Results

During initial data collection it was observed that vocal responses starting with the letter E did not stop the electronic counter on many occasions. Thus, the reaction times to the $M = 4$ ensemble figures were artifactually slow as recorded. Therefore, all data for that ensemble were excluded from further consideration although the subjects continued to view the $M = 4$ ensemble figures throughout so as to achieve the counterbalancing of conditions originally established to assure a Zipf-type relationship between ensemble size and familiarity.

Figure 1 presents the results of Differential Learning Sessions 2, 3, 5 and 6. Data from the first differential learning session were not considered since the immediately preceding session (the final session of preliminary training) had temporarily violated the Zipf-type relationship, as indicated earlier, and we wished to examine the reaction time-memory ensemble size relationship only after the Zipf relationship was reestablished. Data from Session 4 were excluded as these data reflected the subjects' adjustments to the bonus pay system which

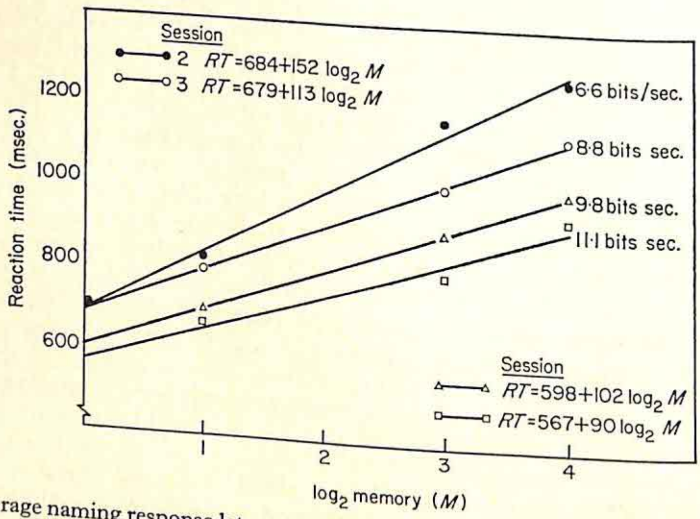


FIG. 1. Average naming response latency as a function of memory ensemble size (M) and practice (sessions).

was introduced at the beginning of that session. Each data point is the mean of 12 medians, one for each subject. Each median was based on 16 responses.

It is apparent from Figure 1 that the first two predictions (see above) were confirmed: As was the case with Oldfield's derivation (1966), there is a linear relationship between vocal naming latency and log memory ensemble size, and the slope constant decreases systematically with practice from 6·6 bits/sec. in Session 2 to 11·2 bits/sec. in Session 6. It is clear, however, that the estimate of 17 bits/sec. derived by Oldfield for meaningful stimulus and response materials still exceeds by far the 11·2 bits/sec. rate of memory retrieval noted here in the sixth session.

It is clear also from Figure 1 that the intercept constants decreased with practice from 684 msec. on Session 2 to 567 msec. on Session 6. Therefore, the subjects became increasingly proficient at stimulus encoding and/or ensemble selection as practice continued from Session 2 through Session 6. An analysis of variance of the data summarized in Figure 1 showed that both memory load and sessions were statistically significant at $P<0{\cdot}01$, while their interaction was significant at $P<0{\cdot}05$.

Figure 2 is a summary of the data from the special test session which were collected on the final session following the sixth differential learning session. The upper function ($E+N$) was determined by following the same procedure as in earlier differential learning sessions: the subject was required to pronounce the full name (ensemble designation plus specific figure name, e.g. "A-24") for each of the figures. The lower function (N/E) was determined in the special test wherein the experimenter announced the ensemble designation just prior to the occurrence of each figure, and so the subjects emitted only the specific figure name, e.g. "24."

The fitted equations are listed in Figure 2. There is no statistically significant difference in slope for these two functions ($P>0{\cdot}05$), but the full-name function

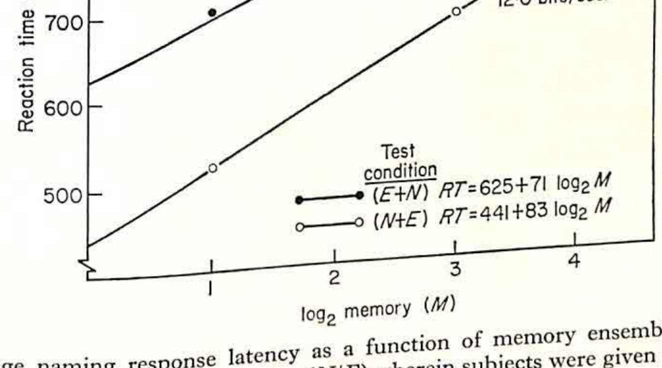


FIG. 2. Average naming response latency as a function of memory ensemble size (M) for a normal testing ($E+N$) and for special testing (N/E) wherein subjects were given advance information on ensemble designation.

$(E + N)$ is significantly inferior overall ($P < 0{\cdot}01$) to the name-given-ensemble function (N/E). Thus, memory retrieval time was the same in both tests, but as predicted, the intercepts differ significantly. We may infer, then, that ensemble selection *per se* required approximately 184 msec. under normal testing on this final session. Once selected, memory retrieval from ensemble occurred at a rate of approximately 13 bits/sec. (the reciprocal of the average of the two slope constants in Fig. 2.)

Discussion

These results confirm Oldfield (1966) in his derivation of the relationship between naming response latency and memory ensemble size. When stimulus-response familiarity is controlled according to a Zipf-type relationship, RT is indeed a linear function of log ensemble size, and the index of ensemble selection time (the intercept constant) does indeed decrease when no ensemble selection step is required. We may conclude with Oldfield, then, that memory for names is organized, in part at least, in terms of familiarity, with the most frequently used names assigned to the smallest memory ensemble, and so on.

These results are still another instance of Hick's Law (1952) relating *RT* to the number of "perceptual choices" (Crossman, 1955) in human information processing. This log relationship was originally interpreted as an indication that central processing occurs in a sequential fashion with dichotomous memory search for the appropriate response. Others (Neisser, 1967) have suggested that a parallel processing mode is more appropriate, and indeed Rapoport (1959) has shown that a log relationship can arise in a parallel as well as in a serial mode of operation. Nevertheless, direct tests of the parallel vs. serial processing question do favour somewhat the serial mode (Egeth, 1966; Nickerson, 1967). Therefore, the present authors favour the following interpretation: stimulus information is encoded, a decision is made based on this information as to which of several memory ensembles is to be searched, that search follows in a dichotomous fashion until a match occurs between encoded and memorial representations, and from this match a response is selected and executed.

While our data support Oldfield's interpretation that part of the intercept constant is due to ensemble selection time, this does cloud somewhat another interpretation that the intercept constant reflects only encoding and decoding times (Sternberg, 1967). If the slope constant of the *RT*-memory load function reflects central, long-term memory retrieval and comparison times, then why is not memory ensemble selection time included in that slope constant? After all, it is a step associated with memory processing. Since the above data indicate otherwise, it follows that selection of a memory ensemble must occur after the initial encoding stage but before the memory search and comparison stages of information processing. This in turn suggests that Sperling (1967) may well be on the proper tack in proposing the presence of a "recognition buffer-memory" between initial iconic storage and long-term storage. Sperling proposed the recognition buffer-memory component specifically for his short-term memory task, and he suggested that it simply selects and stores programs for a later rehearsal of input information (encoded letters). In addition, the present results indicate that such a buffer can respond to input familiarity so as to select an appropriate long-term memory ensemble for further search.

We were sponsored in part by the National Science Foundation through grant GN 534.1 from the Office of Science Information Service to the Computer and Information Science Research Center, The Ohio State University.

References

CROSSMAN, E. R. F. W. (1955). Entropy and choice time: The effect of frequency unbalance on choice responses. *Q. Jl exp. Psychol.* **5**, 41–51.

EGETH, H. E. (1966). Parallel versus serial processes in multidimensional stimulus discrimination. *Percept. Psychophysics*, **1**, 245–52.

HICK, W. E. (1952). On the rate of gain of information. *Q. Jl exp. Psychol.* **4**, 11–26.

NEISSER, U. (1967). *Cognitive Psychology*. New York: Appleton-Century-Crofts.

NICKERSON, R. S. (1967). Categorization time with categories defined by disjunctions and conjunctions of stimulus attributes. *J. exp. Psychol.* **73**, 211–19.

OLDFIELD, R. C. (1966). Things, words and the brain. *Q. Jl exp. Psychol.* **18**, 340–53

OLDFIELD, R. C. and WINGFIELD, A. (1965). Response latencies in naming objects. *Q. Jl exp. Psychol.* **17**, 273–81.

RAPOPORT, A. (1959). A study of disjunctive reaction times. *Behav. Sci.* **4**, 299–315.

SPERLING, G. (1967). Successive approximations to a model for short term memory. In SANDERS, A. F. (Ed.), *Attention and Performance*, pp. 285–92. Amsterdam: North-Holland Publishing Co.

STERNBERG, S. (1967). Two operations in character recognition: Some evidence from reaction-time measurements. *Percept. Psychophysics*, **2**, 45–53.

VANDERPLAS, J. M., SANDERSON, W. A. and VANDERPLAS, J. N. (1965). Statistical and associational characteristics of 1100 random shapes. *Percep. mot. Skills* **21**, 414.

ZIPF, G. K. (1935). *The Psychology of Language*. Boston: Houghton-Mifflin.

Received 10 *February* 1969

Q. Jl exp. Psychol. (1969) **21**, 192–202

BOOK REVIEWS

Fundamentals of Skill. By A. T. Welford. London: Methuen, 1967. Pp. 426. £3 10s.

Dr. Welford spent more than 20 years in the Psychological Laboratory in Cambridge as Lecturer and formerly as Director of the Nuffield Unit for Research into Problems of Ageing, and now he has recorded his experience and understanding of both pure and applied psychological problems in this book. For my part, as an undergraduate I listened enthusiastically to his lectures; I helped at his practical classes while a research student; and more recently assisted him in the production of an advanced course on human experimental psychology. I am therefore very pleased to have the benefits of his teaching and research recorded in detail.

The book is exactly what its title suggests, an analysis of the psychological components of skilled performance, as studied both in the laboratory and in the field. The scope of the book is very wide and its coverage equally comprehensive; the bibliography lists over 1000 references. Principally it consists of a review, in varying detail, and a synthesis of experiments relating to a number of topics in human psychology which have proved of increasing importance in recent years. The underlying ideas are derived more from the mathematical theory of communication than anywhere else and are applied to reaction times (extensively), control of movement, decision-making, both short and long-term memory, and to changes in long-term performance. Dr. Welford is at his best in the intuitive discussion of skilled performance, showing an authority and clarity of expression that come only from long first-hand acquaintance with the original research. The breadth of his comprehension renders the book, in parts, invaluable. The book is weakest in its treatment of information and decision theory, so perhaps this aspect should have the most careful evaluation.

In his introduction Dr. Welford acknowledges a conceptual debt to three books, all published between 1948 and 1950, which have had a profound influence on psychological thinking: he mentions Wiener's *Cybernetics*, Shannon and Weaver's *The Mathematical Theory of Communication* and Wald's *Statistical Decision Functions*. However, the author has been very free in his use and application of certain important ideas, which in their original form are precise mathematical notions, and his dependence on the books he cites is more distant than the reader might expect. An instance appears in the introduction (p. 16). Welford describes the characteristic disjointed performance of a human subject in a tracking task without preview, and says: "We are thus led to view the human mechanisms mediating between sensory input and motor output as a *communication channel of limited capacity* and reaction time as a potentially valuable measure of this capacity." However, on the next page we learn that, if the subject is allowed to see sufficient of the track ahead, his performance is much smoother and more competent. Consequently performance deteriorates in the absence of preview, not because the task presents more information than the communication system can handle, but because the information is presented too late relative to the need for corrective action. Reaction times therefore relate, not to channel capacity, but to the minimal neural conduction time.

The chapters on signal-detection and choice-reaction times have both good and bad aspects. The good points are well exemplified by the comment of an undergraduate (with a mathematical education as incomplete as Dr. Welford's) that the account of signal-detection theory in this book was the only one he could understand. On the other hand, the author goes on to speculate what the decision axis might represent; unfortunately the decision axis in the signal-detection model is not sufficiently well defined operationally for this to be a meaningful question. This part of the book would certainly have been improved by expert criticism while it was in draft. In the chapter on choice-reaction times a fair variety of recent models are discussed briefly and in intuitive fashion. Rather greater space is given to the Communication Theory model, which receives too uncritical a presentation. The treatment of choice-reaction times given here has the advantage of mathematical simplicity, but more accurate and more informative accounts are available elsewhere.

The most unsatisfactory chapter is the one on "Single-channel operation," and this is because Dr. Welford, personally, has had the lion's share in the development of this topic, and his judgement is correspondingly fallible. He is able to write with great authority, in a histriographic sense, and it is interesting to discover how present ideas about central intermittency developed. But he also recounts in disproportionate detail material that is freely available in other recent writings, and which, research-wise, is not of particularly high quality. The trouble is that these ideas have never been spelled out in quantitative detail (this exercise is not attempted in the book), and the analyses of experimental data, which provide the supporting evidence, are generally not appropriate to the precision of the original measurements. Dr. Welford's treatment here compares but poorly with that of, say, R. T. Ollman (in the *Journal of Mathematical Psychology*, 1968), who looked at the single-channel idea as a potential explanation of temporal uncertainty effects in simple reaction time. Ollman showed a great deal more precision and understanding than Welford, and concluded that "the predictions seem to be fairly seriously wrong."

In succeeding chapters Dr. Welford deals with other aspects of human performance, less precise, and therefore less amenable to explicit formulation. These later topics he treats with sensitivity and understanding, and displays his most winning accomplishment, namely, the ability to organize his material within a simple conceptual framework. One may or may not approve his choice of framework, but it certainly makes a wide spectrum of experimentation easy to comprehend and renders the text very readable. From time to time the magic word "information" reappears, generally signalling a temporary disturbance of the balanced understanding and clarity and, above all, honest common sense, which characterize the latter part of the book. In general, I think it difficult to acquire, at second hand a nice understanding of skill, of long-term performance, and of their interactions with the personality of the individual; but Dr. Welford has done a very great deal to help the reader on his way.

This book will be of some use in undergraduate teaching, particularly as collateral reading to a course of lectures; but after the first three chapters, which contain material covered better and more accurately elsewhere, it tends to be a little too detailed, even for third-year students. The book will be very useful as a review and bibliography on skill and human performance for years to come. The coverage of recently reported work is excellent and must have cost much labour in keeping up with the published literature. True, the book will date in this respect, but its review capacity, as of autumn 1967, and its provision of simple conceptual frameworks will be of continuing value for many years.

D. R. J. Laming

Attention in Discrimination Learning. By Elijah Lovejoy. London: Holden-Day. 1968. Pp. xii + 133. £4 7s.

Before starting to praise this monograph, I should declare my own bias. Lovejoy describes a formal version of a two process model of discrimination learning in animals: this model stems from ideas first put forward by Lashley and Krechevsky and more recently developed and extended by Lawrence, Terrace and House, Mackintosh and myself. I am therefore strongly predisposed in favour of the type of theorizing in which Lovejoy indulges.

The book starts with an exceptionally clear discussion of whether animals when trained on a simultaneous discrimination learn to select stimuli or to select responses. For example, an animal being trained to choose black and avoid white, could learn either always to approach black, or to go to the left when black is to the left of white and to go to the right when black is to the right of white. Lovejoy concludes that animals tend to learn approach-avoidance habits, rather than to select positional responses on the basis of the stimulus configuration.

He next discusses the question of whether animals' behaviour is controlled by more than one cue at a time. He presents arguments on both sides of the case but concludes that at least for the purpose of model building we can assume that within a trial animals are controlled by only one cue except in special cases such as conditional discrimination learning. This assumption seems to me rather implausible when one considers the complexity of perception. How many cues do we use when we recognize a human face and how many cues does a rat use to decide that it is in a jumping stand? Lovejoy describes a new experiment of his own on this problem but, despite the ingenuity of the experiment, the results as he admits are not completely convincing.

The most important section of the book is his description of his own formalized version of a two process model of discrimination learning and the results of a computer simulation involving the running of stat-rats. The model is elegant and ingenious and the simulation proves unequivocally that some of the predictions others have derived from loose verbal models of selective attention really can be generated by a formal model. In particular, his stat-rats exhibited position habits very similar to those exhibited by real rats, they displayed the overtraining reversal effect when the relevant cue was not too salient and they showed no ORE when the relevant cue was the most salient cue; they exhibited negative correlations between the amount learned about two relevant cues, and when more than one cue was relevant, breadth of learning was greater under partial reinforcement than under consistent reinforcement.

It is true that, compared with many formal models of the learning process, Lovejoy's has a large number of free parameters (nine when there are three cues and one additional parameter for each additional cue). On the other hand, the range of phenomena to which the model is applicable appears to be very much wider than in the case of any previously published formal model and once the parameters are fixed to account for the results of a given experiment, the model can be tested by discovering whether the results of many different experiments can be obtained with the same parameter settings. There are a great many other phenomena to which the model should be applicable and it is to be hoped that Lovejoy will continue his simulation to explore in detail the strengths (and possible weaknesses) of his model.

One unusual feature of the book is that Lovejoy's Fortran simulation programme is reproduced in full in addition to 20 pages of notes upon it. Having regard to the notorious difficulty of following other people's programmes, this seems a little unnecessary, particularly as the remainder of the text makes it very clear what the programme does. Zeaman and House produced the first formal model of selective attention and some discussion of the differences and similarities between their model and Lovejoy's would have been useful.

The book is very easy to read. The arguments are delightfully clear and terse and the description of the model is unusually easy to follow. The main part of the monograph (excluding the programme) is very short (75 pages) and could be read with profit by undergraduates interested in discovering how one goes about such modelling as well as by research workers interested in following up Lovejoy's stimulating and original ideas. It is a pity that the price/length ratio (about one shilling a page) is so high.

N. S. SUTHERLAND

Eye Movements and Vision. By Alfred L. Yarbus. Translated by Basil Haigh. Translation Editor, Lorrin A. Riggs. New York: Plenum Press. 1967. Pp. xiii + 222. £7 10s.

Tantalizing references to the work of the Russian physicist A. L. Yarbus have been appearing in English-language journals for the last decade. A translation of his monograph is very welcome. It describes his experiments on the recording of all types of eye movements and on stabilised images.

The technical care with which the experiments were conducted is impressive. There is a detailed and helpful account of the construction of the "caps" which, fixed to the eye, are the basis of all the work. After reading this, and the fascinating and coherent description of the experiments on stabilized images, one emerges convinced that the caps adhere to the eye as firmly as the author implies. There remains a puzzling conflict between Yarbus's claim that stabilized images of constant luminance always "become and remain an empty field within 1 to 3 sec.," and the recurrent reappearances found by all other workers. Riggs has recently shown that the residual movements of the stabilized images in his system are probably too small to be visually effective. It would be interesting to know more about the optical quality of the stabilized images in different laboratories.

The readability of the book is evidence of the translator's skill. However, the reader does have to take care in places. For instance, it seemed to me that in Chapter II "disappearing" has sometimes to be read as "being invisible" rather than, as one expects, "becoming invisible."

As Yarbus's bibliography shows, it would be possible to make a collection of papers from the Western literature dealing with almost all the topics covered here. Many of the experiments would antedate those of Yarbus. Most would be more detailed and quantitative, and would contain more searching discussions of causation, and explanation. Nevertheless, this book, as well as confirming old findings and presenting new ones, is for many purposes the most convenient account available. It covers concisely a wide range of topics, is readable and stimulating, and contains numerous helpful illustrations, some in colour.

The range which Yarbus has covered in his own experiments is particularly striking. Riggs writes in his preface, "In the field of Psychology one would perhaps need to go back to the nonsense-syllable experiments of Ebbinghaus to find an experimenter so enthusiastically exploiting a new experimental technique." I am tempted to produce some tit-bits to entice future readers: Under what circumstances does a black-white bipartite field appear uniformly bright red (though with its outer edge visible), and then bright blue? Why can we estimate the length of a spiral as a multiple of the length of a small straight line shown at its centre only if we are allowed to make voluntary eye movements, but not if we have to fixate, even though the spiral remains clearly visible during the fixation? This book should not be thought of as a research monograph of interest only to specialists, but should in spite of its price, be available to all students of perception, from retinal physiologists to cognitive psychologists.

PAUL WHITTLE

Abnormal Hypnotic Phenomena: a Survey of Nineteenth Century Cases. Vol. I, France; Vol. II, Belgium, the Netherlands, Germany and Scandinavia: Vol. III, Russia and Poland, Italy, Spain, Portugal and Latin America; Vol. IV, The United States of America and Great Britain. Edited by Eric J. Dingwall. London: Churchill. 1967–8. Pp. viii + 328 + viii + 256 + viii + 216 + viii + 174. 50s. each volume, or 160s. the set.

In his General Introduction to this series of volumes the editor, Dr. E. J. Dingwall, states that their object is "to fill a gap in the literature of hypnotism in a number of countries." The gap concerned is the one left by the failure of most writers on hypnotism to deal with the somewhat numerous reports of "paranormal phenomena" (such as thought-transmission, eyeless sight, clairvoyance and "suggestion at a distance") which figured prominently in the early history of the subject. There is no doubt that such a gap exists, and no doubt that the volumes under review go a long way towards filling it. They do not, indeed, in any real sense constitute a history; they are precisely what they purport to be—a survey of cases, with appropriate background information and critical comment. Within that limit they are in considerable part successful, and some of the contributions—especially Dr. Dingwall's own—seem to me to strike a high level of scholarship and critical acumen. Anyone who may aspire to writing a full and balanced history of hypnotism, and in particular to understanding the intimate relationship between mesmeric phenomena in the nineteenth century and the genesis of certain religious movements, will find these studies essential reading.

The material surveyed is at times of more than historic interest. Most of the contributors are extremely guarded on the issue of whether or not any of the "paranormal" occurrences described could have been genuine. The great majority of the cases cited seem simply to illustrate varieties of human credulity and of experimental error. But a few—for example those of Adolphe Didier, the famous "clairvoyant," of Khovrin's clairvoyant subject Miss D., and of 'Mme Léonie' with whom Janet and Gibet conducted experiments on hypnosis at a distance—are very curious, and at the least suggest that there may be leakages of information from experimenter to subject along channels more subtle than any which workers on "experimenter affects" have so far uncovered. Some of these nineteenth century case reports bear, it is perhaps worth noting, a strong resemblance to ones currently emanating from the Soviet Union.

This unique survey of "fringe" cases raises other interesting issues of various kinds; and at £8 the set of volumes cannot be regarded as overpriced.

Alan Gauld

Research and Experiment in Stuttering. By H. R. Beech and Fay Fransella. International Series of Monographs in Experimental Psychology. Vol. 6. Oxford: Pergamon Press. Pp. xi + 224. 70s.

To attempt to review even "certain issues illustrative of problem areas in the field" in a literature as diverse as that on stuttering is courageous. Even these issues range from the effect of sugar placebos, carbon dioxide and tranquillizers through self-concepts and goal setting to operant conditioning and delayed auditory feedback in a literature of some 400 references. Unfortunately the sum of all this disparate research is more an object lesson in experimental procedure than a corpus of knowledge on stuttering. This is realized by the authors, who do not attempt to draw more than a few simple conclusions at the end of each chapter.

As in a number of areas of speech research, the central problem is one of definition. The vast majority of studies reported here treat stutterers as a homogeneous group to be compared with a group of controls. Only once do we get a mention of the dangers of this when we are told that the variation within the stutterers was greater than that between the stutterers and the controls. The authors rightly assert that there is no evidence to suppose that stuttering is a unitary disorder.

Perhaps the most surprising lacuna in the research is the almost total lack of phonetic analysis of stuttered speech or of the words most likely to be stuttered. The one study which has considered this has shown that initial rather than final words tend to be stuttered, and consonants rather than vowels, moreover there was a large subjects by consonants interaction. It is perhaps at this level that basic definition must start.

Some of the wilder behaviourist theories of stuttering are given rather an easy time. One theory, which argues that stuttering persists because it is the activity immediately preceding a release from stuttering, escapes unscathed. Nor is the experimental evidence as critically viewed as it might be. It is thought of "great interest" that normal subjects can produce speech disfluency to turn off an electric shock. This surely has as little relevance to the plight of the natural stutterer as the operant conditioning of an arm movement has to Parkinsonism.

The faults of the book derive largely from the research material reviewed rather than from the reviewing itself. The authors have performed a service in bringing together the more experimental research on a little understood disability.

C. J. Darwin

Developments in Applied Psycholinguistics Research. Edited by Sheldon Rosenberg and James Koplin. New York: Collier MacMillan, 1968. Pp. vii + 311. 84s.

This collection of papers, concerned primarily with language pathology, represents the completion of a project that began with the publication of *Directions in Psycholinguistics* (ed. Rosenberg) in 1965.

The best papers are very good indeed. For instance, Blanton's outstanding paper on deafness contains an incisive discussion on the relation between language and cognition, followed by his own experiments on grammatical and syntactic processes in deaf and hearing subjects. He is not afraid to speculate, suggesting that since the deaf child is expected to learn the surface structure of adult English, it would be desirable to teach an English that conforms most closely to the deep structure. In another excellent paper Goodglass identifies a variable which facilitates speech production in agrammatic aphasics, but inhibits it in "fluent" aphasics who have no specific grammatical impairment. Cromwell and Dokecki propose an ingenious "disattention hypothesis" to account for schizophrenic speech. They postulate that the schizophrenics' difficulty in rote learning is solely due to the inability to withdraw attention from distracting stimuli. In contrast, Spradlin's contribution on the language of the retarded is severly practical. He warns us of what he takes to be the unfortunate limitations of negative reinforcers in the shaping of language—"the young child whose hand is caught in a fence may be more likely to cry than to emit an appropriate verbal response that can be reinforced by releasing the hand . . . if a program of aversive control is established for a child, the child will very likely attempt to avoid the adult involved in training." Rosenberg summarizes the papers and expounds the critical differences between the two main competing theories: the "innate language-acquisition system" and Skinnerian learning theory. They are not resolved: the gulf between them is as deep as that between rationalism and empiricism in philosophy, with which it has obvious connections.

In spite of one or two indifferent papers, this is a valuable book which should be read, not only by psycholinguists, but by clinical psychologists concerned with language disorders.

P. C. Wason

Experience and Capacity. Proceedings of the 4th International Interdisciplinary Conference on Learning, Remembering and Forgetting. Vol. IV. Edited by Daniel P. Kimble. New York: N.Y. Academy of Sciences. 1968. Pp. 245.

This book has as its contributors a number of distinguished behavioural scientists. The potential of such a gathering seems to have been dissipated, however, by the form in which the conference is presented. The book appears to be a direct transcription of the taped conference proceedings. This gives the text a disconnected character which is emphasized by the extremely short individual contributions of about 1000 words. This book seems best described as being three written discussion periods, and as such contains all the inherent difficulties of free discussion, especially that of the main topic being lost amidst points which are peripheral.

The first subject discussed is the relationship between active movement and knowledge of visual direction. The main discussions revolve around possible models which might help explain visually determined reaching and the explanation for the observed results of visual deprivation in both kitten and monkey.

The second topic is concerned with the effect of restricted and other experimental rearing procedures on subsequent learning and perceptual discrimination, and the third with the differing effects of hormones or deprivation of hormones at various stages in the development of the organism, and the subsequent effects of such treatment.

There is a good deal of interest in the book as a whole, but because no speaker presents his research or his views in any orderly manner, extracting information presents extreme difficulties; this is made even more trying by the complete absence of an index.

Brian Craske

Foundations of Experimental Research. By Robert Plutchik. New York and London: Harper and Row (Harper's Experimental Psychology Series). 1968. Pp. 290. 37s.

Experimental Psychology: A Methodological Appraoch. By F. J. McGuigan. New Jersey: Prentice-Hall. 1968. (2nd Edn). Pp. 400. 79s. 6d.

Both of these books are concerned with methodology and the design of experiments, yet in style and content they could hardly differ more widely. Plutchik's paperback volume is a lively and highly readable introduction to the subject, pitched at the level of the able undergraduate or beginning graduate student, while presupposing little statistical or mathematical background. The theme of the book is experimentation as a decision-taking process and successive chapters discuss the problems and techniques of decision-taking at different levels or stages of research—decisions concerning definitions of concepts (including an elegant critique of operationism), observation, types of experiment, sampling, bias and control, experimental design, the treatment of functional data, measurement and scaling, and concepts of causality and inference. There is one beautifully lucid chapter on basic statistical concepts, but the book is not concerned with computational methods as such, and in no sense attempts to be a statistics text. In the same vein, just one chapter is spent in discussing conventional experimental design methods, using largely qualitative analyses. His arguments are illustrated from a refreshingly wide range of disciplines, and the discussion of methodology is never allowed to become an end in itself but is constantly related to the historical intellectual process from which it derives. There is a rather conventional chapter on psychophysics with—from the particular angle of this book—a surprisingly sketchy treatment of signal detection theory. The book somewhat tails off at the end with a short chapter on instrumentation in research. But these are minor defects in a book which is sometimes original, always lively, never pompous or obscure.

By contrast, McGuigan's text, aimed explicitly at an elementary undergraduate level, is heavily pedestrian. The core of the book consists of seven chapters (200 pages) on experimental design and statistical methods, treated in very considerable detail—including, for example, quite elaborate factorial designs and a detailed presentation of Duncan's Range-Test—and employing an integrated treatment of both "design" and computational procedures which should be useful. This central section of the book is preceded and followed by chapters which "explore such questions as the nature of science" and "the logical bases of experimental inference." The whole is written in a banal, didactic, rather patronizing style guaranteed to deter all but the most doggedly persistent, and illustrated throughout by irritatingly trivial examples to match. Thus in a footnote on page five McGuigan informs us that experiments relating to intervening variables of any kind will not be discussed in his book "since it is unlikely that your elementary work will involve hypotheses of such an abstract nature." In effect this means that the problems of experimental design are discussed solely in relation to experiments of an exploratory, descriptive, or garbage collection variety, and "hypothesis" is used as equivalent to "prediction." The systematic interrelation of experiment and theory, the derivation of crucial predictions, in a word problems of research *strategy* are simply not considered. Anything more fatally misleading (and sterile) in an introduction to scientific methodology, more especially in one which also purports to provide an introduction to the philosophy of science, this reviewer finds difficult to imagine, albeit this book is by no means the first of its kind.

D. A. ALLPORT

Olfaction and Odours. An osphresiological essay. By W. McCartney. Berlin, Heidelberg and New York: Springer-Verlag. 1968. Pp. viii + 249. 8vo. D.M. 36, $9.00.
Odour Description and Odour Classification: A Multidisciplinary Examination. By R. Harper, E. G. Bate Smith and D. G. Land. London: J. & A. Churchill Ltd. 1968. Pp. viii + 191. 38s.

If the sense of smell claims little space in the modern textbook, this is not for lack of a large experimental, practical and speculative literature. The past 20 years, particularly with the development of molecular physics, chromatographic techniques and the like have considerably added to this corpus and sometimes raised the hope of a fundamental breakthrough. But for the most part of it remains a warren largely ramifying into dead ends. Dr. McCartney takes us on a discursive and not always well-conducted tour of this terrain, the full extent of which will probably not have been appreciated even by those familiar with the more central themes. His book is by turns fascinating and tedious but should be read by everyone interested in the field. It will also gratify those who like to contemplate the failures as well as the successes of the scientific approach.

A good example of a field in which practical, experimental and speculative elements are mixed—not to say mixed-up—and one with which many readers will be unfamiliar is provided by tracker dogs, a topic to which Dr. McCartney devotes over 60 pages. This is a truly astonishing story in which the development of efficient training of canine and human skills directed to a practical end is interspersed with near-mystical or -magical convictions, endless disputation about what it is the dog actually does and would-be scientific investigation which, if often painstaking, seems virtually never water-tight. After some 70 years of work in the field and laboratory, a number of massive monographs, dozens of papers and a World Congress of Kynologists, it still remains uncertain just what the olfactory sensibility or discriminative power of the dog is, or in what way it makes use of the sense of smell in the tracking and searching tasks which, seemingly, it carries out to the satisfaction of its employers.

Dogs, however, form only a substantial opening course to a meal of great variety and sometimes bewildering order. Dr. McCartney touches on birds and fish, the "savage" and the blind, wine tasters and perfumers as well as on such curious byways as the smell of minerals, in particular that of two flints knocked together, which, it seems, is attributable to adsorbed organic substances. He gives some account, too, of earlier and more recent work on the anatomy and physiology of smell and passes in review physico-chemical theories of the mode of action of the end-organ. His handling of all this is not devoid of critical judgement. But it seems a pity that his criticisms are not directed to any convergent clarification of the issues. The whole subject seems if anything, more confused at the end than it was at the beginning. Two points which underlie and partially explain the present confusion, however, stand out with clarity. The first is the extraordinary sensitivity which the human nose can display in certain conditions—a sensitivity which even the most refined of present-day methods of physico-chemical analysis seems unable to match. The second, related, question is that of the character and role of impurities in odiferous substances. In some instances these, occurring in quantities so small as to defy removal, seem to play an entirely disproportionate part in affecting the nature and strength of the odour of substances. The whole field of research remains open to the impact of fresh techniques.

The second book, by a small group of workers who between them have biochemical, psychological and botanical interests, is devoted to a far narrower field. It provides a useful summary and discussion of previously constructed systems of nomenclature and classification and draws together earlier and more recent work in a form which is perhaps not available in any other single source. The outstanding interest of the volume lies, however, in the emphasis put on the psychological processes of verbal classification and of memory factors in the perception and identification of smells. It is, perhaps, a little disappointing that the authors hardly do more than air these aspects generally and do not produce any fresh scheme of description or classification. Their own empirical research is only briefly referred to at the beginning of the book. It is difficult to judge the ultimate value of the present summary and discussion of previous work until these new experimental results are forthcoming.

R. C. OLDFIELD

The Discrimination Process and Development. By B. J. Fellows. Oxford: Pergamon Press. 1968. Pp. 218. £3 15s.

The first half of this book is devoted to the problem of how children learn visual discriminations. The author reviews some of the evidence on the role of external orienting responses and on the part played by mechanisms of selective attention. In the remainder of the book, he considers the difficulties young children have in discriminating between mirror images and between different rotations of the same shape. The influence of development on both processes is reviewed.

The tone of the book is distinctly amiable. A large number of theories are mentioned with approval and many experiments are cited. Unfortunately, the author has been so wide in his choice of material that he is only able to deal with any of it in rather a cursory way and one feels he has not got down to the business of sorting out the wheat from the chaff. His own theory of discrimination learning is an amalgam of almost all previous theories, but it does not seem to be rigorous enough to allow clear predictions to be drawn. It is never, for example, stated whether changes in orienting responses or in internal mediating responses depend upon rewards and punishments or whether they are learned in some other way.

Perhaps the most interesting part of the book is the account of the various pragmatic techniques that have been devised to speed up discrimination learning in young children. Although it is still not possible to give a rigorous explanation of why some of these techniques are more successful than others, at least they provide food for thought.

Despite the fact that the author does not attempt to cut very deep (or possibly because of this fact), the book is clearly written and easy to read. It should help to stimulate interest in an important field and provides a useful guide to a wide range of literature.

N. S. SUTHERLAND

Human Color Perception. By Joseph J. Sheppard, Jr. New York: Elsevier Publishing Company. 1968. Pp. xvii + 192. £6.

This is "an attempt to provide an introduction to the critical study of the experimental foundation of human color perception" for the would-be research worker. The author, an engineer by training, rather surprisingly disclaims the expertise "to select the material necessary to cover the essentials of the subject," but hopes to give "both the flavour and some of the meat in each of the several subdivisions of color perception." These aims are fulfilled, and Sheppard's modesty in accurately describing the limitations of his book disarms criticism. The selection of material, although patchy, is sometimes refreshing, such as in its descriptions of the physical properties of receptors. The writing is clear, and some of the critical comments usefully dispel complacency with the current state of the subject.

PAUL WHITTLE

Analysis of Behavioral Change. Edited by Lawrence Weiskrantz. New York, Evanston and London: Harper and Row. 1968. Pp. 447. $12.50.

In his Introduction, Professor Weiskrantz explains that this book has resulted from his reluctance to accept an invitation to review "psychopharmacology." He felt that such a review could be of merely transitory significance, largely because the inherent problems of analysing changes in behaviour had not recieved adequate attention. A review and discussion of these problems would therefore be more enduring, and also more useful—not only for the analysis of interactions between behaviour and drugs. This book is devoted to that end. It attempts to give the flavour of the problems, and an indication of how they have been tackled, or even how they might be tackled in the future.

The book consists of chapters written by different authors, although six of them are by Weiskrantz himself. Each chapter has a general, fairly arbitrary title. This inevitably leads to a certain amount of repetition in places. On the other hand, most of the chapters are packed with carefully chosen data to illustrate the problems, and many of them (particularly those by Weiskrantz) are replete with insights of interpretation which make the book as a whole unusually stimulating.

The first substantive chapter is on "Positive Reinforcement," by Steiner. This lays most emphasis on the problems of interpreting changes in the efficacy of a positively reinforcing stimulus. This is undoubtedly an important problem, but perhaps one might claim that the effects of such stimuli can be adequately represented only by including an examination of schedule-effects, a topic which is mentioned (by defining a few schedules) but discussed in no great detail. The chapter on "Negative Reinforcement" by Morgan explores the problems of analysing behaviour controlled by "aversive" stimuli, but, again, although there are some thoughtful and original comments, the relative neglect of the importance of schedules of presentation (as shown by recent work) makes exhaustive analysis difficult. A chapter on "Emotion" by Weiskrantz is for this reviewer particularly outstanding, and certainly succeeds in presenting important problems of interpretation forcefully and elegantly. In addition, it suggests much useful and necessary research on this topic. Chapters on "General Activity" (Gross) and "Exploratory Behaviour" (Halliday) provide rigorous reviews; although the relevance of the latter to "behavioural change" might have been made more explicit. The next four chapters, "Learning" by

Sheldon, "Memory" by Weiskrantz, "Discrimination" by Cowey, and "Some Aspects of Attention" by Weiskrantz seem to suffer most from the difficulties of arbitrary fragmentation, for as their titles show, none can be really independent of the others. However, each chapter may be regarded as a useful interpretative essay, each being very firmly based on the exposition of data rather than theories. In particular, the contribution by Cowey presents data from the whole range of the general topic, both learned and unlearned discriminations. In this and the chapter on attention, references to physiological research and to theories of interpretation are made with the apparent intention of *complementing* work on overt behaviour, not replacing it, and this is one of the overall themes of the book which this reviewer retains most strongly. One might claim that work on errorless discrimination learning has more importance in the context of the book's aims than Cowey gives it, and also that the concept of "stimulus control" might be more carefully examined in his chapter and in that on attention. Following these chapters, there is a lengthy, but fascinating, review of "Alteration of Perception and Memory in Man" by Brenda Milner and H.-L. Teuber, which exhibits most adequately the emphasis on controlled experimentation for which these writers are admired. This contribution alone would make a viable book: the complexities of assessing behavioural changes in humans are most tellingly and clearly reviewed, mainly in the context of the effects of brain lesions, although the general implications are always clear. There follow two further chapters on aspects of animal experimentation, "The Neurological Examination of Animals" by Ettlinger, and "Ethological Methods of Observing Behavior" by Bateson. These chapters are less exciting, for although both examine areas of exceptional challenge and interest, they fail to develop their points forcefully, and, indeed, both chapters are relatively short of data. One's reactions are made more marked perhaps, by the contrast with the preceding contribution. The book concludes with two more general chapters by Weiskrantz, the first a useful consideration of "inner workings" of organisms in so far as our appreciation of these helps us to interpret behaviour, and the second, "Some Traps and Pontifications," making some sound general points about the interpretation of drug-effects on behaviour.

The above is intended to provide a brief indication of the contents of the book. There are many constructive points made which deserve discussion, so many that it becomes impossible in a review of this nature.

In the light of recent publications (e.g. "Behavioral Pharmacology" by Thompson and Schuster), one might just query the reluctance of Weiskrantz to examine in more detail the interactions between drugs and behaviour. Moreover, it is not clear whether the data presented in the book *should* be regarded as secondary to the problems of interpretation. Will the general problems discussed survive future empirical advance as Weiskrantz suggests? In fact, one of the most outstanding features of the better chapters seems to be the pertinent exposition of data, and constructive suggestions when more data is needed. Again, it is sometimes difficult to believe that the book is mainly concerned with behavioural *change*. However, this book is to be most warmly welcomed. In particular, it should play a useful part in the development of behavioural science in this country. For example, its derivation of data from operant conditioning experiments is carried out in a non-doctrinaire way, thereby, one hopes, revealing their general significance to any experimenters who may be deterred from this approach by the dogmatism of some of its exponents. The same might be said for the book's physiological and theoretical discussions. Specific criticism is slight indeed in comparison with this general welcome.

Perhaps the best adjective to describe this book is *useful*. My copy has certainly been used a great deal already.

DEREK BLACKMAN

Publications Received

Books

Adjustment and Mental Health. By A. Arkoff. London and New York: McGraw Hill Book Co. 1968. Pp. xvi + 536. 83s. 6d.

Space and Time. By R. Swinburne. London and Toronto: MacMillan Press. 1969. Pp. vii + 319. $9.00.

Deviancy: The Psychology of Being Different. By J. L. Freedman and A. N. Doob. New York: Academic Press 1968. Pp. vii + 158. 75s.

Human Adaptation and Its Failures. By L. Phillips. New York: Academic Press. 1969. Pp. xiii + 271. 98s.

Formal Representation of Human Judgement. Edited by B. Kleinmuntz. London and New York: John Wiley & Sons Inc. 1968. Pp. xii + 273. 75s.

Group Dynamics (Research & Theory). Edited by D. Cartwright and A. Zander London: Tavistock Publications. 1968. Pp. ix + 580. 65s.

Frames of Mind. By L. Hudson. London: Methuen & Co. Ltd. 1968. Pp. ix + 134. 25s.

Industrial Training Handbook. Edited by J. W. Barber. London: Iliffe Books Ltd. 1968. Pp. 13 + 412. 70s.

The Self in Social Interaction. Edited by C. Gordon and K. J. Gergen. London and New York: John Wiley & Sons Inc. 1968. Pp. xii + 473. 90s.

The Regulation of Behavior. By D. E. Lunzer. Staples Press. 1968. Pp. xvi + 392. 70s.

Introduction to Theoretical Linguistics. By John Lyons. London: Cambridge University Press. 1968. 70s.

Digital Computing. By Richard S. Lehman and Daniel E. Bailey. New York: John Wiley & Sons Inc. 1969. Pp. xxiv + 303. 79s.

Paperbacks

The Human Use of Human Beings. By N. Wiener. London: Sphere Books. 1968. Pp. 186. 6s.

The Beginnings of Modern Psychology. By W. M. O'Neil. Harmondsworth: Penguin Books. 1968. Pp. 153. 13s. 6d.

Disorders of Memory and Learning. By G. A. Talland. Harmondsworth: Penguin Books. 1968. Pp. 176. 35s. 6d.

Assessment in Clinical Psychology. By C. E. Gathercole. Harmondsworth: Penguin Books. 1968. Pp. 174. 14s. 6d.

Psychometric Assessment of the Individual Child. By R. D. Savage. Harmondsworth: Penguin Books. 1968. Pp. 117. 15s. 6d.

The Ecological Perspective on Human Affairs. By H. Sprout and M. Sprout. Princeton University Press. 1965. Pp. xi + 236. $2.95.

The Nature of Emotion. Edited by M. B. Arnold. Harmondsworth: Penguin Books. 1968. Pp. 397. 9s.

Q. Jl exp. Psychol. (1969) **21**, 203–208

EFFECTS OF GRAMMATICALNESS, PRESENTATION RATE, AND MESSAGE LENGTH ON AUDITORY SHORT-TERM MEMORY

DAVID GERVER

Department of Psychology, University of Durham

The aim of this study was to demonstrate the effect of message structure, message length, and presentation rate upon auditory short-term memory. Subjects were asked to reproduce in writing messages at different levels of "grammaticalness" (Coleman, 1965) presented through headphones at 3–4 words/sec. or 6–7 words/sec. There was better recall of the more structured, shorter, and slower messages. The results are discussed in terms of acquisition and encoding of information.

INTRODUCTION

Having to attend to someone who speaks at a very fast rate usually imposes severe strain on the attention and memory mechanisms of the hearer, particularly if the speaker maintains a rapid delivery for a long period of time, and if his subject matter or style is complex. Though the present study does not deal with the problem of attention over long intervals, it is hoped that the results will demonstrate that the structure of the message, the length of the message, and the rate of delivery have a significant effect on the short-term memory performance of the hearer.

In a recent review of the effects of rate of presentation, and message structure on recall of auditory messages (sets of digits or words), Neisser (1967) suggests that rapid presentation will have a detrimental effect on short-term memory when rate of presentation exceeds the rate at which items can be encoded. Aaronson (1965) and Norman (1966) have similarly suggested that, when rate of presentation exceeds rate of acquisition (i.e. encoding), performance on short-term memory tasks will suffer because of an initial failure to acquire a stimulus rather than through decay of an insufficiently acquired trace. Neisser supports his argument by citing the work of Yntema, Wozencraft and Klem (1964), who found that immediate recall of digits was impaired by rapid presentation (8-10 words/sec.), and suggest that "during the presentation of a slow list the subject performs some sort of process that makes it easier for him to recall the list a moment later, but there is not enough time for him to perform this process when the list is presented rapidly." Yntema *et al.* employed a "slow" rate of 2-4 words/sec. According to Neisser, the effect of rapid presentation is not so pronounced when the input has already been grouped, or structured, for the subject. If this is the case, then one would expect that the linguistic structure of the message would be an important factor for the listener faced with a very fast speaker.

In coping with speech input the encoding procedures discussed by Neisser would, then, involve the utilization of the syntactic and semantic properties of language which have recently been shown to play such an important role in the perception of, and memory for, language. The more highly structured the speech input, the more readily it should be acquired (and hence recalled) when rapidly presented since the subject will not have time to fit relatively unstructured input to the syntactic and semantic frameworks so necessary for correct perception and recall.

Message length may also affect the listener's ability to assimilate information at very fast rates of presentation. Yntema *et al.* (1964) also found better recall for shorter lists than for longer lists at fast rates of presentation, but not at slower presentation rates. Using rates of 3 words/sec. and 1 word every 3 sec., Laughery and Pincus (1966) found a similar interaction between presentation rate and sequence length in a task involving immediate recall of letters.

The aim of the present study, then, is to examine the effect of rate of presentation, language structure and message length upon auditory short-term memory. The material used in the study was developed by Coleman (1965) in order to demonstrate that a generative grammer could produce word sets which native speakers of a language could rank in terms of their "grammaticalness." The degree of agreement between the grammar and subjects' linguistic intuition was further reflected in the positive relationship between level of grammaticalness and ease with which subjects learned the word sets.

Method

Materials

The 40 word-sets (10 scales at each of 4 levels of "grammaticalness") devised by Coleman (1965) were recorded without intonation, and in random order, on tape by the experimenter at a fast rate. Word-sets of "level 1 grammaticalness" were generated by Coleman by drawing words at random from samples of English prose; level 2, by restricting the choice of words by introducing some of the rules of a generative grammer; levels 3 and 4, by implementing progressively more of such rules. Five of the scales contained word-sets 9 words long, one scale had word-sets 7 words long, three scales were 6 words long, and one consisted of sets 3 words long.

The recorded tape was then processed by means of an Eltro Tempophon. Briefly, the Tempophon compresses or expands the speech signal by varying the speed of rotation of one cylindrical rotating head containing four equally spaced replay heads. The gap between each head is equivalent to 30 msec. This form of processing does not significantly alter either the pitch or the overall quality of the signal, provided that maximum levels of compression or expansion are not required. A control study employing a 9 subjects × 9 levels of compression/expansion × 9 sentence list design showed losses of intelligibility of only 7 per cent at 100 per cent expansion, and 5 per cent at 100 per cent compression of the speech signal.

For the purposes of this study the original recording was processed with equal degrees of compression and expansion to produce two stimulus tapes. The faster rate was approx. 6–7 words/sec.; the slower rate approx. 3–4 words/sec. With the normal rate of talking estimated at approx. 2·75 words/sec. (Carroll, 1964), speech at 3–4 words /sec. can be said to sound like rather fast normal speech, whereas at 6–7 words/sec. speech is very fast and sounds rather clipped and abrupt, but still highly intelligible.

Subjects

The subjects were 20 undergraduate psychology students.

Design

For the purposes of this study the 9-word sequences were designated as "longer", and those of 7 words or less as "shorter." A split-plot design was employed for which the subjects were divided into two groups, one of which heard the longer word-sets at the faster rate, and the shorter at the slower rate. The second group received the reverse order of presentation. Half of the subjects received the slower rate of presentation first, and the other half the faster rate first.

Procedure

The recordings were played back to subjects individually through Koss Pro-4 headphones, from a Ferrograph tape recorder. Subjects were told that they would hear more or less meaningful sets of words, and that they were to write down as much as they could recall of each word-set as soon as they had heard it. Each word set was preceded by the word "Ready" spoken at a normal rate. There was a gap of 30 sec. between word-sets.

Results

Table I presents a summary of the results, which were recorded in terms of per cent words correctly recalled in the correct order.

TABLE I

Means and standard deviations of words correctly recalled in correct order

Rate of presentation	Fast					Slow				
Level of grammaticalness	1	2	3	4	X̄	1	2	3	4	X̄
Short	70·5	61·8	61·7	78·6	68·2	86·8	86·8	82·1	96·3	88
σ	9·1	10·5	7·4	10·7	11·8	5·9	8·3	8·6	2·7	8·6
Long	38·7	47·9	60·1	70·7	54·4	66·6	76·2	83·1	93·4	79·8
σ	6·7	8·8	8·6	6·6	14·4	11·5	12·1	10·1	5·1	14·7
X̄	54·5	54·9	60·9	74·7	61·3	76·7	81·5	82·6	94·9	83·6
σ	17·8	11·9	8·1	9·7	14·9	13·6	11·7	9·5	5·0	12·1

An analysis of variance was carried out, and, in order to ensure homogeneity of variance, the percentage scores were transformed to arc-sin values. The analysis of variance showed the following effects (1) The effect of presentation rate was significant beyond the 0·001 level ($F = 151·22$; $d.f. = 1, 18$); significantly better recall occurring with the slower than with the faster presentation rate.

(2) The effect of level of grammaticalness was also significant beyond the 0·001 level ($F = 52·81$; $d.f. = 3, 54$); significantly more being recalled at the higher than at the lower levels of grammaticalness. There was no significant interaction between presentation rate and level of grammaticalness. Scheffe's test of multiple comparisons on the transformed treatment sums (Edwards, 1960) showed differences between presentation rates at all levels of grammaticalness to be significant beyond the 0·001 level.

(3) The significant rate of presentation × groups interaction ($F = 8·88$; $d.f. = 1, 18$; $P < 0·01$), and the significant presentation rate × level × groups interaction

($F = 37{\cdot}18$; $d.f. = 3, 54$; $P < 0{:}001$) both reflect the effect of length of word set. In order to partial out the effects of length, presentation rate, and the level of grammaticalness Scheffe's test was again carried out on transformed treatment sums. The decrement in recall for longer as opposed to shorter word-sets was significant beyond the 0·001 level ($A' = 3{\cdot}0230$, Crit. A $= 1{\cdot}6308$). At the faster rate the difference between longer and shorter word-sets was significant beyond the 0·001 level ($A' = 1{\cdot}7749$; Crit. A $= 1{\cdot}6308$), while at the slower rate of presentation this difference was significant at the 0·01 level ($A' = 1{\cdot}2693$; Crit. A $= 1{\cdot}1826$). At both rates of presentation the differences between recall of longer and shorter word-sets were only significant at level 1 of grammaticalness. For the faster rate $A = 2{\cdot}1939$ ($P < 0{\cdot}001$), and for the slower rate $A = 1{\cdot}2157$ ($P < 0{\cdot}01$). These results are illustrated in Figure 1.

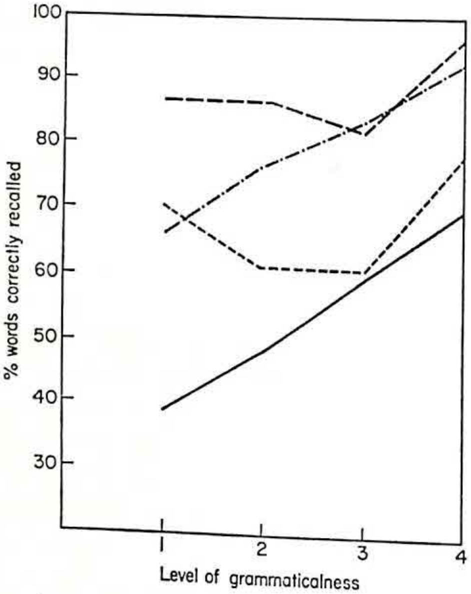


FIGURE 1. Mean percentage of words correctly recalled at each level of grammaticalness: — — —, short slow; —·—·—, long slow; — — — —, short fast; —, long fast.

An analysis of order errors in recall showed that only 25 order errors were committed by 12 out of the 20 subjects. Even though the total number of such errors is small, it can be seen from the distribution of these errors shown in Table II that the majority of order errors occurred in the recall of longer, slower word-sets at lower grammatical levels. Other than omissions, there were no other types of recall error.

TABLE II

Order errors in recall

	Fast				Slow			
Level	1	2	3	4	1	2	3	4
Short	1	1	0	0	1	0	0	0
Long	3	0	1	0	11	7	0	0

Discussion

The results of this experiment clearly support the conclusions of the many previous studies which have shown that syntactic structure facilitates retention and recall of language. The greater the constraints imposed by the rules of the grammar from which the messages were constructed, the greater the ease of encoding and recall. Inspection of Figure 1, however, reveals that this does not appear to hold for shorter word-sets. Though there is a monotonic increasing relationship between grammaticalness and recall for longer word-sets, the relationship is non-monotonic for word-sets shorter by two words or more. The non-linearity of the curves for shorter word-sets suggests perhaps that Coleman's rules for generating word-sets are more likely to provide cues to structure in longer than in shorter word-sets.

The overall effect of rate of presentation on subjects' performance supports the assumption that some time is required for encoding and storage, but the lack of significant interaction between presentation rate and level of grammaticalness does not appear to support the hypothesis that faster presentation rate would impair recall of relatively unstructured material. Inspection of the data shows, however, that when message length is taken into account, there was poorer recall of longer messages both at the faster presentation rate and at lower levels of grammaticalness. Since there is no significant difference in recall of longer and shorter messages at higher grammatical levels, the longer messages are within the memory span. It would seem, then, that the longer messages only exceed the span when there is neither time nor structure enough with which to encode them. In other words, longer messages are more effectively encoded when the structure is immediately available, and unstructured longer messages are more effectively encoded when enough time is available.

These conclusions agree with Neisser's interpretation of previous work, and with Pollack (1952) who suggested that subjects would show improvement in recall at slower presentation rates because the extra time between items could be used for encoding and organizing the stimulus information. Pollack suggested that the time immediately after the presentation of a stimulus is a critical one because of the processes which relate input to previously stored information. In the case of speech, then, "previously stored information" can be interpreted in terms of the syntactic and semantic rules of language.

The effects of message length, structure and presentation rate on recall are consistent with the function of the type of two stage perceptual process discussed by Aaronson (1967). Input may first be held in a large capacity fast-decay buffer storage prior to processing by a limited capacity secondary storage system. When the rate of input to the primary system is both rapid and unstructured there appears to be a greater likelihood that, in the case of longer messages, traces could decay before processing at the secondary stage. Though there are perhaps not sufficient data in the present study to permit any but the most tentative conclusions, and even allowing for the greater possibility of order errors in recall of longer messages, the analysis of order errors does show that a greater proportion of such errors occurred at the slower presentation rate, at lower levels of grammaticalness, and in longer messages. As would be expected, structure provides cues to order in storage and recall, cues which are more important in longer sequences, which, when unstructured, approach the limits of short-term memory capacity. Decreasing the time available for processing the input, and/or decreasing the amount of the signal available for perception (by increasing presentation rate through speech compression) has led to reduced accuracy in recall in terms of overall amount recalled, but not in terms of order errors. This result is to be expected if the fast presentation rate exceeds acquisition rate, thereby reducing the amount of material available for recall. The subject, therefore, has less material in store for restructuring on recall.

References

AARONSON, D. (1965). Perception and immediate recall of auditory sequences. Paper presented to the Eastern Psych. Assoc. 1965.

AARONSON, D. (1967). Temporal factors in perception and short-term memory. *Psychol. Bull.* **67**, 130–44.

CARROLL, J. B. (1964). *Language and Thought*. Englewood Cliffs. Prentice-Hall Inc.

COLEMAN, E. B. (1965). Responses to a scale of grammaticalness. *J. verb. Learn. verb. Behav.* **4**, 521–7.

EDWARDS, A. L. (1960). *Experimental Design in Psychological Research*. New York: Holt, Rinehart and Winston.

LAUGHERY, K. R. and PINKUS, A. L. (1966). Short-term memory: Effects of acoustic similarity, presentation rate and presentation mode. *Psychon. Sci.* **6**, 285–6.

NEISSER, U. (1967). *Cognitive Psychology*. New York: Appleton-Century-Crofts.

NORMAN, D. A. (1966). Acquisition and retention in short-term memory. *J. exp. Psychol.* **41**, 329–35.

POLLACK, I. (1952). The effect of rate of presentation of information. *Human Res. Lab. memo rep.* No. 25.

YNTEMA, D. B., WOZENCRAFT, F. T. and KLEM, L. (1964). Immediate serial recall of digits presented at very high rates. Paper presented to the Psychonomic Society, Niagara Falls, Ontario.

Received 14 *February* 1969

Q. Jl exp. Psych. (1969) **21,** 209–213

ERRORS AND ERROR DETECTION IN TYPING

L. H. SHAFFER AND JANE HARDWICK
Department of Psychology, University of Exeter

The errors and error detections made by qualified and trainee typists were examined, when they were given a passage of prose that included pseudo-words and words made up of random consonants. It was shown that trainees make more and detect relatively fewer errors. For all typists errors are made more often per letter in long or degraded words and are relatively less likely to be detected.

Introduction

In two previous papers (Shaffer and Hardwick, 1968, 1969, unpublished results) an analysis was made of the performance of skilled touch typists given systematically varied texts to transcribe. In all the experiments reported the typist was instructed not to correct her errors, since time to completion was the primary measure of performance. Analysis was thus based upon speed and error scores, with no opportunity to examine error detection.

The mechanism of error detection is of particular interest in touch typing because, in the nature of the skill, the typist is usually looking neither at the type she has just produced, nor at the letter in the text she has just copied but perhaps five to seven letters ahead of this in the text (Butsch, 1932).

In the present experiment the typist was instructed to stop as soon as she detected an error and retype the word in which it occurred. It was originally hoped to examine speed of detection, but this was soon answered: most often the word was abandoned as soon as the error was made and if abandoned later it was simplest to assume that the typist had failed to override habit. However a classification of errors based upon subject and text variables provides an informative picture that expands our knowledge of the transcription process.

Method

Subjects

The touch typists were taken from two populations differing in typing experience: these were 51 qualified typists from departments in the university and 33 trainees who had just completed a secretarial course at the technical college.

University subjects were tested one at a time, but the trainees were tested together.

Apparatus

Each typist used the typewriter on which she normally worked or was trained.

The text was a passage of prose typed in double spacing on a foolscap sheet. A number of words in the passage, randomly located and covering different word lengths, were degraded such that half of these were reduced to nonsense words having syllabic structure and the other half to random strings of consonants.

Instruction

The typist was told to transcribe the text without altering its layout and it was pointed out that the text would include nonsense and random words. She was told to type as quickly as possible and if she felt she had made an error she should abandon the word immediately and retype it. (Sufficient margin was given in the text to accommodate word repeats without altering line layout.)

Results

Total errors

Altogether 1576 errors were recorded, 817 by qualified typists and 759 by trainees. The distribution of errors per subject shows that significantly more were made by the trainees, χ^2 ($d.f. = 1$) $= 52$, $P < 0{\cdot}001$.

The words in the text were classified as Long (>8 letters), Medium (6 to 8 letters) or Short (<6 letters); they were also classified as Real, Nonsense or Random (see above). There were 1426 letters of text and their distribution in classes of word is shown in Table I.

TABLE I

Frequencies of errors in text words classified according to length and structure

	Real	Nonsense	Random	
Long	296 (23)	162 (11)	375 (11)	833 (45)
Medium	130 (14)	104 (8)	227 (8)	461 (30)
Short	111 (17)	64 (4)	107 (4)	282 (25)
	537 (54)	330 (23)	709 (23)	1576

The distribution of letters in each word class is shown as percentage in brackets.

Also shown in Table I are the frequencies of errors in each word class. Comparing these frequencies against their expectation according to letter distribution then the following are shown: (1) errors increase significantly with word length, χ^2 ($d.f. = 2$) $= 51$, $P < 0{\cdot}001$; (2) they increase significantly between Real, Nonsense and Random words, χ^2 ($d.f. = 2$) $= 469$, $P < 0{\cdot}001$; and (3) on the combined classification of word length and structure errors appear far above expectation in Long Random and Medium Random words, particularly the former, and below expectation in Short Real words, χ^2 ($d.f. = 4$) $= 498$, $P < 0{\cdot}001$.

Comparing errors made by the different groups of subjects in different word classes then it is shown that whereas their distributions of errors were similar for word length, they differed significantly for word structure, χ^2 ($d.f. = 2$) $= 9{\cdot}2$,

$P < 0{\cdot}05$, the qualified group making relatively fewer errors than the trainee group on Real words. The error scores are shown in Table II.

TABLE II

Error frequencies by qualified and trainee typists as a function of word structure

	Real	Nonsense	Random
Qualified	250	182	385
Trainee	287	148	324

Error detections

Breaking down errors into those detected and those undetected then overall about one-third of errors were detected but these were distributed differently in the two subject groups, relatively more detections occurring in the qualified group, $\chi^2 (d.f. = 1) = 21, P < 0{\cdot}001$. These scores are shown in Table III. In addition there were 32 false positives—abandonment of words correctly typed—but their occurrence per subject is too low to warrant analysis.

TABLE III

Frequencies of detected and undetected errors made by qualified and trainee typists

	Detected	Undetected
Qualified	314 (38)	503
Trainee	209 (28)	550

Per cent detections are shown in brackets.

Breaking down detected and undetected errors in different word classes we obtain the results in Table IV. The left table exhibits a significant relationship between detection and word length, χ^2 $(d.f. = 2) = 6{\cdot}9$, $P < 0{\cdot}05$, there being relatively fewer detections in Long words. The right table exhibits a much stronger relationship between detection and word structure; χ^2 $(d.f. = 2) = 33{\cdot}6$, $P < 0{\cdot}001$, the probability of detection declining steadily between Real, Nonsense and Random words.

TABLE IV

Frequencies of detected and undetected errors as a function of word class

	Detected	Undetected
Long	252 (30)	581
Medium	168 (36)	293
Short	103 (36)	179

	Detected	Undetected
Real	217 (40)	320
Nonsense	124 (38)	206
Random	182 (26)	527

Per cent detections are shown in brackets.

In case detection probability varies over the conjoint classification of word length and structure a χ^2 test for a three-way contingent relationship was carried out. This failed to reject the null hypothesis at the 5 per cent level.

There is reason to expect that qualified and trainee typists should differ in detection performance as a function of word structure. A χ^2 test for a three-way contingent relationship between subject groups, detection and word structure failed to reject the null hypothesis at the 5 per cent level. However the largest single term in the χ^2 showed precisely that qualified typists tend to make relatively fewer undetected errors in Real words than trainees.

Discussion

Reducing word structure or, to a lesser extent, increasing word length leads to an increase in typing errors and to a decrease in detection probability. These text factors were also shown in the earlier studies (*ibid.*) to lead to a decrease in typing speed. The trainee would presumably be slower, on average, than the qualified typist; she is in any case more prone to error and detects fewer of her errors. The effect of typing experience on error performance is magnified with Real words and we are inclined to believe that on these the qualified typist detects relatively more of her errors as well as make relatively few of them, although only the latter is observed to be statistically significant. In fact the detection probabilities obtained on Real, Nonsense and Random words were, 0·50, 0·43 and 0·29 for qualified and 0·32, 0·30 and 0·22 for trainee typists.

In the earlier paper it was stated as hypothesis that a skilled typist reads the text in linguistic fragments of such a size that each fragment can be (a) readily decomposed into the letter string that spells it, (b) held as an ordered sequence of letters in an output store and (c) translated letter by letter into response, with some overall acceptable risk of error. Loss of speed occurs if the readability of text reduces such that it can only be read in improvised fragments of suboptimal length, thus slowing transfer into and out of store. The increase of errors in Nonsense or Long words may be a consequence of having to convert improvised units of text into letter strings. Or it may be that the typist attempts to maintain her speed—in the experimental situation—by accepting a higher risk of error.

If transcription and error detection are respectively primary and secondary information processes, sharing the same processing resources (Bahrick, Noble and Fitts, 1954), then any factor that increases the load of the first, such as difficulty of text encoding, should reduce the output of the second. Conversely if practice simplifies the primary or both processes then error detection should improve. Thus the gross results of detection are predictable. One can however pursue the analysis further to examine new questions.

We have previously shown (1968) that errors may be classified as Response, if a key adjacent to the correct one is pressed; Context, if two letters are transposed; reading, if a letter that looks or sounds like the original is typed; or Omission, if a letter is left out. These classes cover most of the errors observed. It was hoped to analyse detection in each error class, but unfortunately if a typist abandons a word as soon as it is in error then we cannot distinguish subsequently whether the error was an omission or would have been a transposition. The results as far as they can be classified are shown in Table V. However the ambiguities are resolved; it appears that error detection occurs in each class.

TABLE V

Frequencies of detected and undetected errors in different categories of error

	Detected	Undetected
Response	200	520
Reading	18	74
Omission	18	149
Context	87	205
Omission or context	147	—
Miscellaneous	33	105

Let us suppose that the basis of error detection is a comparison of feedback information of the response, probably kinesthetic, with some derivation of the stimulus (cf. Rabbitt, 1968). Since the typist is reading ahead in the text then detection of Reading and Omission errors implies either that she uses regressive eye-movements to re-cover the text or that she stores a copy of each text fragment until it is typed. Butsch (1932) recorded regressive eye-movement but these were few compared with forward saccades. Again, if Context errors occur as a loss of ordinal position in the output store then detection of these implies access to a copy of the input to this store. A similar argument applies to Response errors and the translation of letters-in-store into responses. Whether each stage of transcription retains a copy of its input for error detection or whether an image of the text caters for all levels of feedback comparison is a question we cannot yet answer. Whatever the nature of the image it must allow serial scanning in phase with the response output.

We gratefully acknowledge the financial support of the Science Research Council for this research, which is part of a study on serial skills. We also thank Miss E. Holding, in the Department of Business Studies at Exeter Technical College, who arranged for us to test the trainee typists.

References

BAHRICK, H. P., NOBLE, M. and FITTS, P. M. (1954). Extra-task performance as a measure of learning a primary task. *J. exp. Psychol.* **48**, 298–302.

BUTSCH, R. L. C. (1932). Eye movements and the eye-hand span in typewriting. *J. educ. Psychol.* **23**, 104–21.

RABBITT, P. M. R. (1968). Three kinds of error-signalling responses in a serial choice task. *Q. Jl exp. Psychol.* **20**, 179–88.

SHAFFER, L. H. and HARDWICK, J. (1968). Typing performance as a function of text. *Q. Jl exp. Psychol.* **20**, 360–69.

Received 17 *February* 1969

Q. Jl exp. Psychol. (1969) **21**, 214–224

RECOGNITION VS. RECALL: STORAGE OR RETRIEVAL DIFFERENCES?

RICHARD D. FREUND, JOHN W. BRELSFORD, JR.† AND RICHARD C. ATKINSON
Stanford University, Stanford, California 94305

Differences between recognition and recall performance may be due to differences in storage processes, differences in retrieval processes, or some combination of both. An attempt was made to determine which process was critical by withholding information, at the time of study of a stimulus-response pair, about how that item was to be tested on its next presentation. It was found that differences between recognition and recall did not depend upon whether or not the subject knew, at time of study, the mode of test to be employed. These results were interpreted as support for the assertion that, in this particular task, differences in retrieval processes were sufficient to account for differences in recognition and recall. It was found that both the direction and magnitude of the recognition-recall difference depended upon the guessing correction employed.

Introduction

Differences in recognition and recall performance may be attributed to differences in storage processes, differences in retrieval processes, or some combination of both (Atkinson and Shiffrin, 1968). The present experiment is an attempt to determine whether recognition-recall differences may be observed in the absence of differential storage processes.

Consider the learning of a list of paired associates, where the set of response alternatives is well known to the subject. The study of a stimulus-response pair will be assumed to generate some input $\mathbf{I}$ to the memory system. Storage processes may be represented by an operator $\mathbf{S}$ applied to this input. The information stored in memory, denoted by $\mathbf{I}'$, is therefore a function $\mathbf{S}$ of the input:

$$\mathbf{I}' = \mathbf{S}(\mathbf{I}).$$

At the time of test a retrieval operator $\mathbf{R}$ is applied to this stored information, which results in output $\mathbf{O}$:

$$\mathbf{O} = \mathbf{R}(\mathbf{I}').$$

In short, the output is a function $\mathbf{R}$ of what is stored, which is in turn a function $\mathbf{S}$ of the input:

$$\mathbf{O} = \mathbf{R}(\mathbf{S}(\mathbf{I})).$$

---

† Present address: Yale University, New Haven, Conn., U.S.A.

This functional representation of memory permits a logical analysis of the various operations involved in recognition and recall. Typically the subject, throughout an experimental session, is aware of which mode of test is to be employed. Thus there is ample opportunity for utilization of different storage operations and different retrieval operations. If recognition and recall differ in storage processes, then what is crucial to the differential storage is knowledge of the mode of test to be employed. Without this knowledge there cannot be differential storage, and hence any performance differences may be attributed to retrieval processes alone.

In the present experiment, a list of paired associates was learned using an anticipation method. The procedure varied from the typical one, however, for any item could be tested either by recall or recognition on any trial, i.e. when an item was studied on trial n the learner did not know how it was to be tested on trial $n + 1$. Thus there was no opportunity for storage to take place in one way in anticipation of a recall test, and in another way in anticipation of a recognition test. Any observed differences between recognition and recall performance therefore would be attributable to different retrieval processes. Two control lists were also used in the experiment, one learned employing recall tests throughout, and the other, recognition tests. For these lists the subject knew at all times how an item was to be tested.

Method

A within-subjects design was employed, with each subject tested under all experimental conditions in each experimental session. The two independent variables were: (1) method of testing—recall or recognition; and (2) whether or not the subject knew, when studying a given stimulus-response pair, how that pair was to be tested.

Stimuli and responses

The stimuli were 432 three-letter nonsense syllables (Consonant-Vowel-Consonant pronounceable trigrams) of approximately 20 to 50 per cent association values. They were randomly selected from the Glaze list in Appendix A of Underwood and Schulz (1960). Eight different lists of 54 stimuli each were constructed. Each subject received a different list on each session. Throughout the experiment the responses were the digits 1 to 9.

Apparatus

Programming, generation of stimuli, and response recording were controlled by an on-line PDP-1 computer. Stimuli were electronically generated and displayed on the face of a cathode-ray tube (CRT). Responses were made on an electric typewriter located directly beneath the lower edge of the CRT. Each subject was seated in an individual sound-proofed experimental booth for each of his sessions.

Procedure

An anticipatory paired-associate learning procedure was used. The subjects were eleven Stanford students who received two dollars per session. Three to five practice sessions on other lists were completed for each subject to insure his understanding of the instructions. The experimental sessions were then initiated. No subject had fewer than five nor more than seven experimental sessions, and a total of 65-subject sessions were completed. For each subject-session one of the 54-item lists was selected at random. This list was subdivided into three 18-item lists. Responses were assigned randomly within each 18-item list, with the restriction that each response was assigned to exactly two stimuli.

The three 18-item lists can be distinguished by the type of retention test employed:

Recall (Re): One list of 18 items was always tested under recall conditions. On a test the stimulus would appear on the CRT, and the subject's task was to recall the response which was paired with the stimulus on prior study trials of the item.

Recognition (Ro): A second list of 18 items was always tested under recognition conditions. On a test the stimulus would appear on the CRT, along with the correct response and one incorrect (distractor) response. These response alternatives appeared to the right of the stimulus, one above the other, with the correct choice randomly located in the two positions. The subject's task was to choose the correct response for a given item from the two response alternatives displayed. The distractor was selected randomly on each trial for each recognition test.

Mixed: The third list of 18 items represented the crucial experimental manipulation. An item in this list could be tested either by recall or by recognition, the choice determined randomly at the time of test. Subjects were instructed that if the stimulus member appeared alone on the CRT they were to try to recall the correct response. If the stimulus member appeared with two response alternatives (displayed exactly as in the Ro condition) they were to try to recognize the correct response. Thus, a given item in this list could be tested on one trial for recall, but on the next trial could be tested either for recall or for recognition. This point was emphasized in the instructions. Those items from this "mixed" list which on any particular trial happened to be tested by recall shall be denoted as Recall Mixed (ReM) items and those tested by recognition as Recognition Mixed (RoM) items.

A "trial" consisted of one presentation of all 54 items. Within a trial the order of events was as follows: (1) A stimulus was selected from one of the three 18-item lists, e.g. from the Recall list, and was presented for test. The subject was given 3 sec. to respond; if he did not respond in this time period a message appeared on the CRT encouraging him to go faster. Next, the correct answer for that stimulus was displayed for 0·75 sec., followed by a 0·50 inter-item interval. (2) A stimulus from one of the remaining two lists, e.g. from the Mixed list, was selected and presented for test. Again the test period duration was 3 sec.,

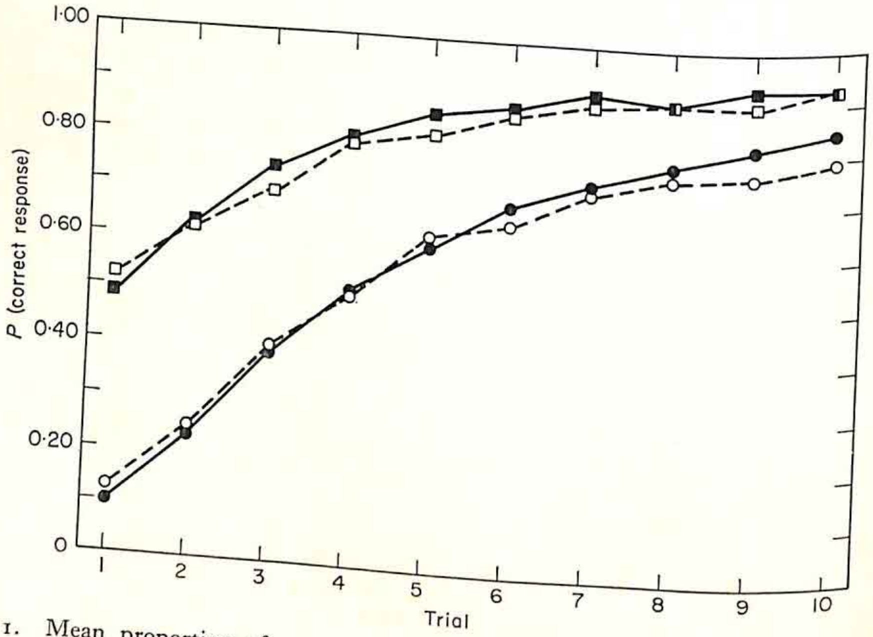


FIGURE 1. Mean proportion of correct responses as a function of trials for all conditions. —■—, Recognition; - - □ - -, recognition mixed; —●—, recall; - - ○ - -, recall mixed.

correct feedback was displayed for 0·75 sec., and inter-item interval was 0·50 sec. (3) A stimulus was selected from the remaining list (here, e.g. Recognition) for test and study just as for the two previous lists. Then this same order of selecting items would be repeated for three more items, and so on, until all 54 items had been tested. This represents trial 1. Each 18-item list was then instantly re-randomized and trial 2 began (without noticeable interruption). Ten such trials were completed in each session by each subject. Each day the subject received a new list of 54 stimulus-response pairs to learn. There are six permutations of the orders of testing of Recall, Recognition and Mixed; these permutations were assigned randomly among the 11 subjects (but remained fixed for each subject over all his sessions).

Crucial to the design is the fact that subjects knew what type of item was being presented at all times. To emphasize the distinction between the three types of items, they were displayed in a different third (top, centre or bottom) of the CRT. In addition, the appropriate label RECALL, RECOGNITION or MIXED appeared in capital letters to the far left of the stimulus (and remained on during the test period) for the three types of items.

The instructions explained the basic paired-associate learning task: learn the correct pairings between the CVC's and the digits. The three types of procedures were described in approximately the same fashion as in this report. Subjects were not told the number of items to be learned. They were instructed to make guesses (from number 1 to 9) if they could not recall the correct response. On recognition tests subjects were required to choose between the two response alternatives displayed on the CRT.

Results

The proportion of correct responses for each condition is plotted in Figure 1, averaged over all subjects and sessions. There are approximately 1400 observations at each point on the Recall (Re) and Recognition (Ro) curves, and 700 observations per point on the Recall Mixed (ReM) and Recognition Mixed (RoM) curves. It is clear that Recognition did not differ from Recognition Mixed, and that Recall did not differ from Recall Mixed. Analysis of variance was performed on the observed proportion of correct responses, summed over the ten trials

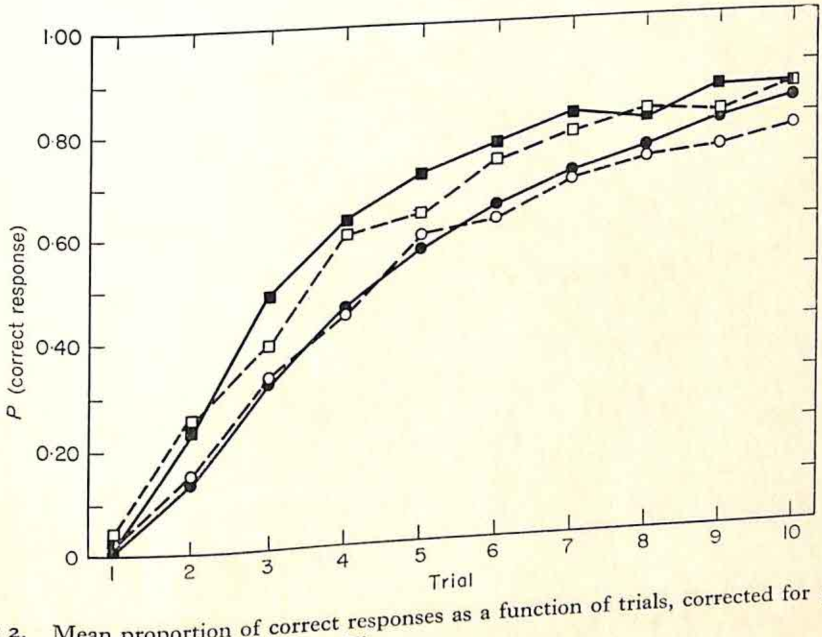


FIGURE 2. Mean proportion of correct responses as a function of trials, corrected for guessing For explanation of symbols, see legend to Fig. 1.

for each subject and condition, using the Newman–Keuls method (Winer, 1962, p. 114). Re vs. ReM and Ro vs. RoM did not approach significance. The four other pairwise comparisons (Re vs. Ro vs. RoM, Ro vs. ReM, ReM vs. RoM) were significant at the 0·01 level.

In order to compare recognition and recall performance some account must be taken of correct responses attributable to chance. A correction procedure frequently used (Hilgard, 1951, p. 556) involves the following transformation:

$$p' = p - \frac{1 - p}{N - 1},$$

where p denotes the observed proportion of correct responses in a given condition, p' the transformed proportion, and N the number of response alternatives (i.e. $N = 2$ for recognition tests, and $N = 9$ for recall tests). One interpretation of this transformation is that the observed proportion of correct responses is a weighted average of those items correctly retrieved from memory and those items correctly guessed.

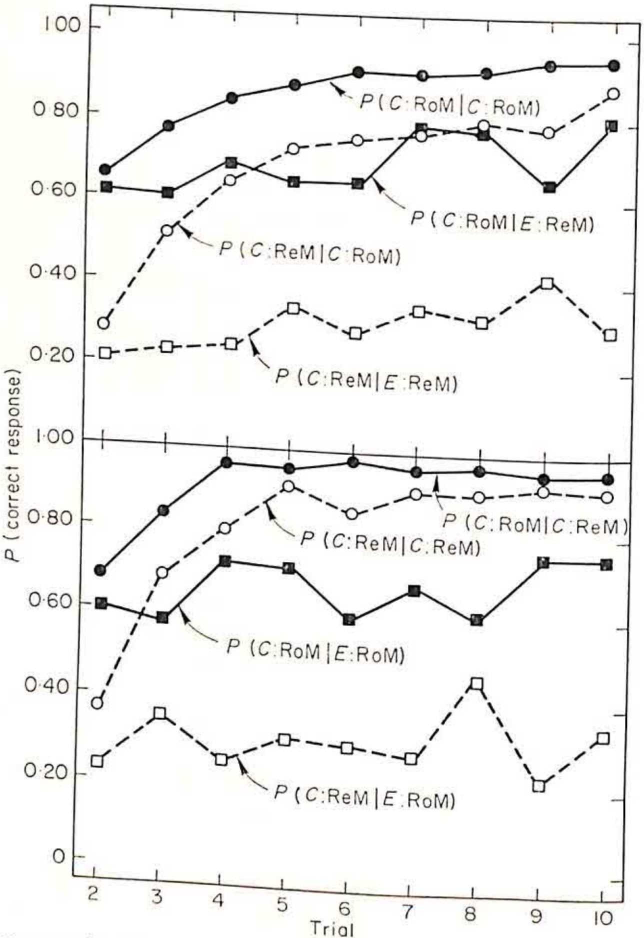


FIGURE 3. Mean proportion of correct responses as a function of trials for the Mixed Condition, conditionalized upon both the type of test and the response given on the preceding trial.

Application of this correction for guessing yielded the data plotted in Figure 2. Because of the correction for guessing all curves start at or near zero proportion correct. The learning of the recognition items appears faster than that of the recall items. The Newman–Keuls method of analysis of variance leads to rejection of the hypothesis that performance in the four conditions was equal ($F = 18{\cdot}48$, d.f. $= 3/30$, $P < 0{\cdot}01$). The six possible paired comparisons were then considered, and again it was found that Re vs. ReM and Ro vs. RoM did not approach significance. All other comparisons were highly significant ($P < 0{\cdot}01$). The reader should be alerted to the fact that these results are very much dependent upon the specific guessing correction employed; we shall return to this point later when we consider an alternative method for correcting for guessing.

An item in the mixed condition could be tested on any trial either by recall or by recognition. Figure 3 presents the proportion of correct responses on trial n for either type of test, conditionalized upon the mode of test on trial n-1 and also upon whether or not the response on trial n-1 was correct or an error. The legend code is $P(C{:}i\,|\,j{:}k)$, where $i =$ mode of test on trial n, $j =$ correct (C) or error (E) on trial n-1, and $k =$ mode of test on trial n-1. For example, in the upper panel of Figure 3 the uppermost curve, labelled $P(C{:}\mathrm{RoM}\,|\,C{:}\mathrm{RoM})$, represents the proportion of correct RoM responses on trial n, given a correct response on the previous RoM test trial for that item.

All curves in Figure 3 were tested for stationarity by Kendall's tau for multiple observations. The hypothesis of a constant (conditional) proportion correct over trials could be rejected for all but two of the curves at the 0·05 level of significance. The two stationary curves are both in the lower panel, namely, $P(C{:}\mathrm{RoM}\,|\,E{:}\mathrm{RoM})$ and $P(C{:}\mathrm{ReM}\,|\,E{:}\mathrm{RoM})$.

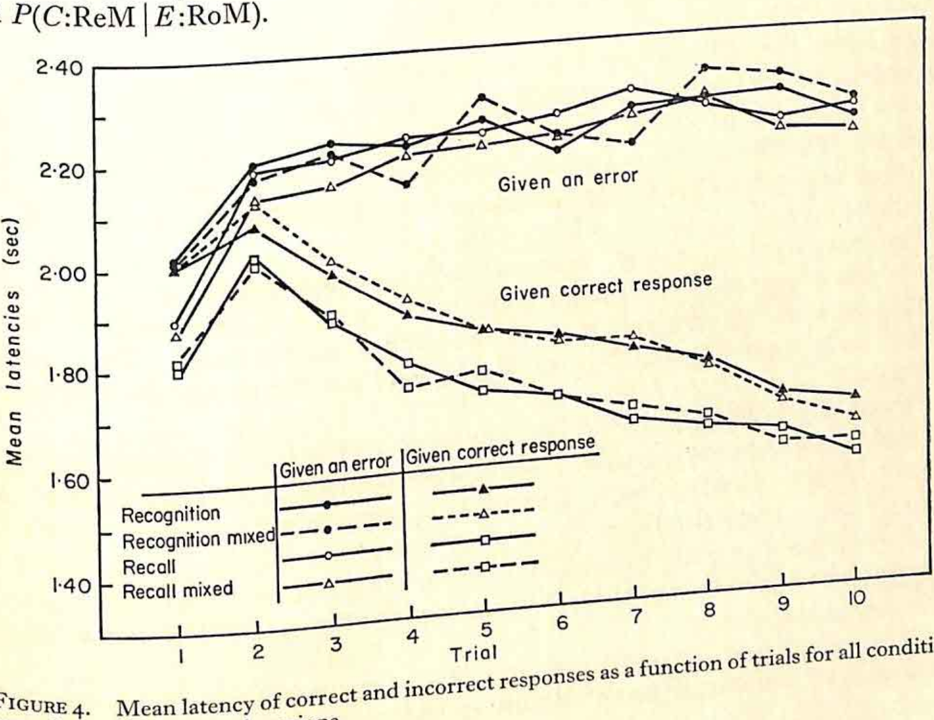


FIGURE 4. Mean latency of correct and incorrect responses as a function of trials for all conditions, averaged over subjects and sessions.

Latencies of responses conditionalized upon whether the response was correct or incorrect are presented in Figure 4. While latencies of incorrect responses remained fairly constant after the first trial (on which all responses were necessarily guesses), the latencies for correct responses showed a steady decrease over the course of the session. It appears that incorrect responses had equal latencies for all conditions, constant at approximately 2·2 sec. For correct responses, the two recall curves decreased together and the two recognition curves decreased together, but the latter were consistently about 0·1 sec. slower at each trial.

Discussion

The results indicate that the assumption of differential storage processes is not necessary to account for the differences between recognition and recall performance. The argument can best be presented by use of the simple equations introduced to relate input and output to the storage and retrieval operations. Recognition and recall tests may logically employ different storage and retrieval operations. Define $\mathbf{S}_{\mathrm{ro}}$ and $\mathbf{R}_{\mathrm{ro}}$ to be the storage and retrieval operators, respectively, for the recognition condition. Similarly, $\mathbf{S}_{\mathrm{re}}$ and $\mathbf{R}_{\mathrm{re}}$ are the storage and retrieval operators for recall conditions. To represent the storage processes in the mixed condition one more operator, $\mathbf{S}_{\mathrm{m}}$, is required. Note that retrieval on the mixed tests utilizes either $\mathbf{R}_{\mathrm{ro}}$ or $\mathbf{R}_{\mathrm{re}}$. The output functions for the four conditions are therefore as follows:

$$\text{Ro:}\quad \mathbf{O} = \mathbf{R}_{\mathrm{ro}}(\mathbf{S}_{\mathrm{ro}}(\mathbf{I})) \qquad (1)$$
$$\text{RoM:}\quad \mathbf{O} = \mathbf{R}_{\mathrm{ro}}(\mathbf{S}_{\mathrm{m}}(\mathbf{I})) \qquad (2)$$
$$\text{ReM:}\quad \mathbf{O} = \mathbf{R}_{\mathrm{re}}(\mathbf{S}_{\mathrm{m}}(\mathbf{I})) \qquad (3)$$
$$\text{Re:}\quad \mathbf{O} = \mathbf{R}_{\mathrm{re}}(\mathbf{S}_{\mathrm{re}}(\mathbf{I})) \qquad (4)$$

Equations (1) and (2) show the relation between performance on the standard (i.e. where the test mode was known) and mixed recognition tests. Since only the storage operator is different, and since there were in fact no performance differences between these two conditions, we conclude that $\mathbf{S}_{\mathrm{ro}} = \mathbf{S}_{\mathrm{m}}$. The coincidence of the standard and mixed recall conditions leads to the conclusion, by inspection of equations (3) and (4), that $\mathbf{S}_{\mathrm{re}} = \mathbf{S}_{\mathrm{m}}$. Thus, all storage operators were identical.

The two mixed conditions, represented by equations (2) and (3), differ only in the retrieval operators applied to the stored information. It was between these conditions that the observed performance differences occurred. The conclusion is that differences between $\mathbf{R}_{\mathrm{ro}}$ and $\mathbf{R}_{\mathrm{re}}$ account for the performance differences obtained.

While the preceding analysis is logically consistent, it could be argued that neither differential storage nor differential retrieval took place in the present experiment. Rather, the superiority of recognition reflects only the greater "value" of identical information. For example, retrieval of the information, "The correct response is a high number," is quite valuable on a recognition test which presents 2 and 8 as response alternatives. Recall performance, given the same bit of partial information, would be inferior.

The existence of such partial information in the present experiment is suggested by the conditional analyses of Figure 3. From trial 4 onward it is clear that correctly recalling a response implied virtually perfect recognition performance on the following trial. Conversely, failure to recognize an item led to near chance performance on the following trial for either test mode.

The above mentioned argument against either storage or retrieval differences clearly employs a more limited use of the term "retrieval" than does the present paper. Shiffrin and Atkinson (1969) have classified retrieval into three primary mechanisms: search, recovery, and response generation. Search is the process of locating the stored information in memory. Recovery is the process by which some or all of the information is made available for the generation of a response. This response generation is considered to comprise all aspects of translating the recovered information into the desired response. These aspects include the development of guessing strategies adopted within the context of particular test requirements. In other words, retrieval (as used here) includes the differential utilization of partial information, and therefore the differences obtained may appropriately be attributed to "retrieval" processes.

The finding that the conditional latencies for errors remain relatively constant while those for correct responses decrease over trials, is the typical result for paired-associate studies of this type (cf. Bower, 1967). The coincidence of the two recognition conditions, and that of the two recall conditions, lends further support to the hypothesis that identical processes took place regardless of whether or not subjects knew how they were to be tested.

More noteworthy is the fact that the correct recognition responses were consistently slower, from trial 2 onward, than those for correct recall responses. This may be explained in part by the following model, which assumes that latencies for responses retrieved correctly from memory have a fixed mean value L, whereas a retrieval failure and subsequent guess has a fixed mean value L'. Thus error responses always have a mean latency L'. Correct responses may occur as a result of a retrieval from memory or a correct guess, and consequently the latencies for correct responses represent a weighting of L and L'.

This model would be expected to have fair success for the following reason: a higher proportion of the correct responses in the recognition conditions were due to correct guessing. The relative weight of L' is therefore higher for recognition, and the latencies for corrects would be expected to be longer. An attempt to derive the quantitative predictions of the latency results, however, from this weighting function in conjunction with the observed probabilities of correct responses at each trial, revealed that this model is not adequate.

On correction of raw scores for chance successes. The success of the present experiment in the separation of storage and retrieval is dependent upon the existence of recognition-recall performance differences. The establishment of such differences raises the problem of correcting raw scores for guessing. The most common transformation is the one employed in this report. An alternative transformation of the raw scores is provided by the theory of signal detectability (TSD). Murdock (1966) has discussed the application of TSD to evaluation of recall and recognition performance. Table II in Elliott (1964, pp. 682–683) allows direct conversion of observed proportion of correct responses to d' as a function of the number of alternative responses.

The raw data of Figure 1 were transformed to d' scores and are replotted in Figure 5. The 9-choice values may be considered as the analogue of the present experiment's recall task if we ignore the fact of the physical absence of the 9 alternatives at time of test. The 9-choice values were obtained by plotting the theoretical 9-choice function on normal-normal co-ordinates, estimating the slope and y-intercept, and applying the linear approximation formula given by Elliott (1964, p. 680).

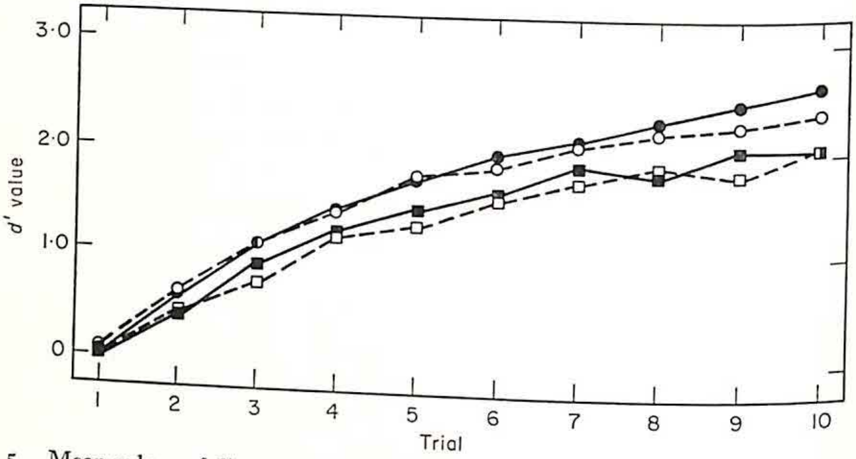


FIGURE 5. Mean values of d' as a function of trials for all conditions, averaged over subjects and sessions. For explanation of symbols, see legend to Figure 1.

While it was expected that the d' transformation might attenuate the superiority of recognition over recall, the completely surprising result obtained was a reversal of the direction of superiority: recall was now superior to recognition. To explain how this could come about, Figure 6 presents the theoretical results of

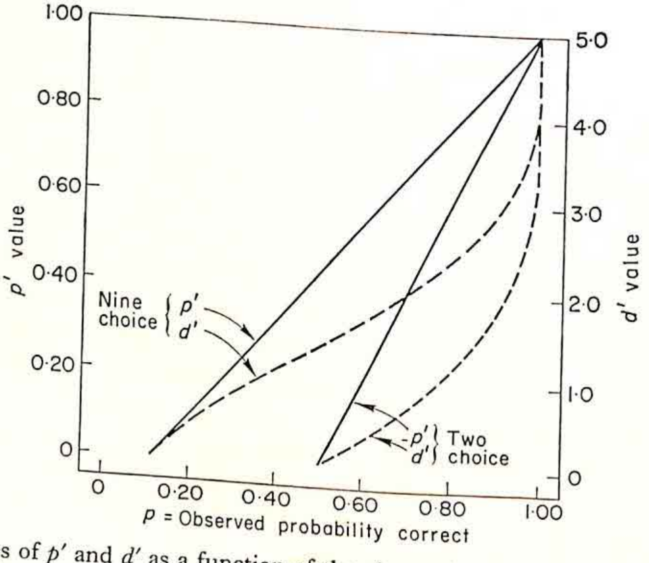


FIGURE 6. Values of p' and d' as a function of the observed probability of a correct response p, for the 2-choice and 9-choice tasks.

transforming 2 and 9 choice data by the standard correction for guessing (p') and TSD (d'). The former transformation yields the linear relation between the observed probability p and p'. The curved lines (to be read against the right-hand ordinate) show the d' value as a function of p.

Define a "reversal" as the event where one transformation of the data makes 2-choice superior to 9-choice and the other transformation yields the opposite ordering. An example of such a reversal occurs when the observed values are $p_2 = 0{\cdot}80$ and $p_9 = 0{\cdot}60$ (the subscripts indicate the number of alternatives). A p' transformation of these scores yields $p_2' = 0{\cdot}60$ which is greater than $p_9' = 0{\cdot}55$, whereas a d' transformation of the same scores yields a $d_2' = 1{\cdot}19$ which is less than $d_9' = 1{\cdot}70$. Hence the same observed scores yield opposite ordering; the p' transformation indicates that recognition is superior to recall, whereas d' suggests the opposite conclusion.

Figure 7 plots the entire space of all possible observations of p_2 and p_9. The two functions plotted may be called iso-probability curves: they are the locus of values of p_2 and p_9 which must be observed to yield $p_2' = p_9'$ (for the linear function) and $d_2' = d_9'$ (for the bowed function). If both 2-choice and 9-choice performance are at chance (i.e. $p_2 = 0{\cdot}50$ and $p_9 = 0{\cdot}111$) then by either transformation the two performances will be judged equal and therefore the two curves intersect; the same is true for the case where $p_2 = p_9 = 1{\cdot}00$.

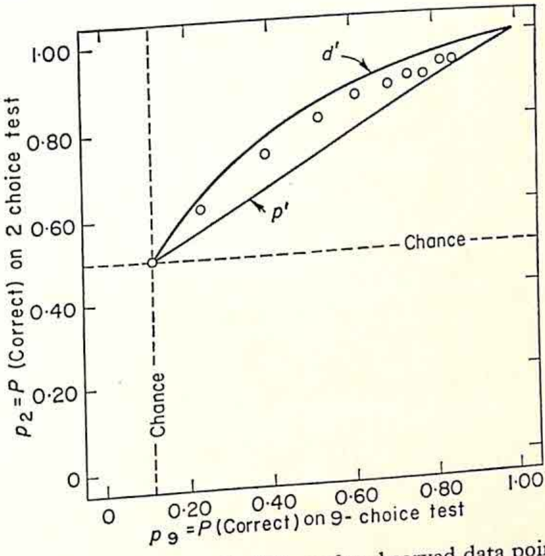


FIGURE 7. Relationship between p' and d'. (When the observed data point falls below the linear function, both the p' and d' corrections indicate that recall is better than recognition; for a data point above the bowed function both yield recognition as superior to recall. In the middle region, the p' and d' corrections yield opposite conclusions.) ○, Data observed on recall and recognition.

Consider now some other points in this space. If the observed data point (p_2, p_9) lies above the linear functions then using the p' transformation we conclude that recognition is better than recall (namely, that $p_2' > p_9'$). However, using the d' transformation, the observed data point (p_2, p_9) must fall above the bowed function for us to reach the same conclusion (namely, that $d_2' > d_9'$). Hence, any data point bounded between the two functions will yield a reversal: the p' correction indicating that recognition is superior to recall, whereas the d' correction yields the opposite conclusion. The data points for the present experiment (the observed values of p_2 and p_9 on each trial for the Re and Ro conditions) are plotted in Figure 7, and as expected, lie in this middle region.

It is clear that the question of whether or not recognition is superior to recall remains meaningless until the correction for guessing is specified and defended. The defence of any such correction rests upon an assumption: namely, that the correction employed is reasonable with respect to the stated theory about the nature of the memory trace. It is maintained, on the basis of the preceding analysis, that special care must be given to the support of such assumptions before recognition and recall can be compared in any meaningful way.

Support for this research was provided by the National Aeronautics and Space Administration, Grant NGR-05-020-244.

References

ATKINSON, R. C. and SHIFFRIN, R. M. (1968). Human memory: A proposed system and its control processes. In SPENCE, K. W. and SPENCE, J. T. (Eds.), *The Psychology of Learning and Motivation: Advances in Research and Theory*, Vol. 2, pp. 89–195. New York: Academic Press.

BOWER, G. H. (1967). A descriptive theory of memory. In KIMBLE, D. P. (Ed.), *The Proceedings of the Second Conference on Learning, Remembering and Forgetting*, pp. 112–85. New York: New York Academy of Sciences.

ELLIOTT, P. B. (1964). Tables of d′. In SWET, J. A. (Ed.), *Signal Detection and Recognition by Human Observers*, pp. 651–84. New York: Wiley.

HILGARD, E. R. (1951). Methods and procedures in the study of learning. In STEVENS, S. S. (Ed.) *Handbook of Experimental Psychology*, pp. 517–67. New York: Wiley.

MURDOCK, B. B. (1966). The criterion problem in short-term memory. *J. exp. Psychol.* **72**, 317–24.

SHIFFRIN, R. M. and ATKINSON, R. C. (1969). Storage and retrieval processes in long-term memory. *Psychol. Rev.* **76**, 179–193.

UNDERWOOD, B. J. and SCHULZ, R. W. (1960). *Meaningfulness and Verbal Learning*. New York: Lippincott.

WINER, B. J. (1962). *Statistical Principles in Experimental Design*. New York: McGraw-Hill.

Received 22 February 1969

Q. Jl exp. Psychol. (1969) **21**, 225–238

SIZE CONSTANCY IN MONKEYS WITH INFEROTEMPORAL LESIONS

N. K. HUMPHREY† AND L. WEISKRANTZ†
Psychological Laboratory, University of Cambridge

In a study of size constancy, monkeys were trained with extended practice to choose the larger of two projected stimulus discs independently of distance. When surgical brain lesions were made following pre-operative training, it was found that posterior parietal lesions caused no deficit on this task while inferotemporal lesions caused a severe breakdown from which three out of four animals showed almost no recovery. Mathematical analysis showed that the pattern of performance after inferotemporal lesions could be described as due to a tendency to oscillate between a "correct" and two "incorrect strategies" and that this tendency was also apparent to at least some extent in the unoperated animals. It is argued that the two incorrect strategies might have resulted from a disturbance of size constancy, such that the animal became unable to use both retinal size and distance information in computing physical size and instead used either retinal size or distance information alone. Parallels are drawn between such an hypothetical perceptual disorder and certain clinical disturbances in man.

Introduction

Removal of the inferotemporal cortex in monkeys creates a mosaic of deficits in the performance of visual tasks which no single theory has been able to piece together. As solutions themselves become fragmented, one simple possibility deserves consideration still. This is that the lesion produces a primary disturbance in the perception of visual form. Such a disturbance, acting to distort visual experience of present events and hence necessarily leaving its mark on associative memory and imagery, could be responsible for many of the deleterious effects on visual behaviour which have been observed.

A lead to the kind of disturbance which might occur can be obtained from human clinical cases. There is a recurrent type of perceptual disorder in man characterized by fluctuating changes in the size, shape and position of objects in the visual fields (see Critchley, 1953; Willanger and Klee, 1966). Patients report that the perceived world undergoes dramatic transformations: objects may grow suddenly large or small, they may move or tilt to one side and their contours may become unstable and undulating. The illusory changes sometimes occur sporadically but are likely to be brought on especially by prolonged fixation (Willanger and Klee, 1966). The disorder in man is associated with several types of central pathology. It can occur in cases of organic brain damage, where the lesion lies in the posterior parieto-temporal or occipital cortex. But very similar symptoms are found in cases of temporal lobe epilepsy (Penfield and Jasper, 1954), the early stages of schizophrenia (Chapman, 1966), recovery from perceptual isolation (Heron, Doane and Scott, 1956), and mescalin intoxication (Mayer-Gross, Slater and Roth, 1960); they can be produced also by direct electrical stimulation of parietal or temporal cortex (Hecaen, Penfield, Bertrand and Malmo, 1956).

---

† Present address: The Institute of Experimental Psychology, 1 South Parks Road, Oxford, England.

The research to be described here looked for evidence in inferotemporal monkeys of a perceptual disorder potentially comparable to such disturbance in man. It was designed in particular to test a hypothesis about the functional origin of the disorder should it be present. This is, that it is due to a failure of perceptual constancy for size and form, i.e. a failure in the process whereby the observational context in which an object is seen—the observer's distance, position of his eyes, angle of regard—is taken into account in assessing the properties of the object from its retinal image.

Methods

Outline of the experiment

Monkeys were trained on a visual task requiring the judgement of size independently of distance. Their performance was compared before and after surgical brain lesions in the hope of detecting some change attributable to a failure of constancy.

The visual task was modelled on that used by Holway and Boring (1941) in their classical study of size constancy in man. Holway and Boring asked their subjects to make size matches between a variable and a standard disc projected on screens at different distances down two long corridors. For the present study, instead of a size match a size discrimination was required, the monkeys being trained to choose the *larger* of two projected discs.

There were two operated groups, the experimental group with bilateral inferotemporal lesions and a control group with bilateral posterior parietal lesions.

Subjects

Seven young experimentally naive rhesus monkeys were subjects. Three, P1-P3, were assigned to the control group and four, IT1-IT4, to the experimental group.

Surgery

Anaesthesia was induced with Nembutal. Subpial cortical excisions were made by aspiration with a fine sucker. At the end of the experiment the brains were perfused with formalin, embedded in paraffin, sectioned at 20 μ and every fifth section stained with thionine. Histological constructions of the lesions are shown in Figure 1.

Apparatus

The stimuli were presented as circular discs of light, back-projected on to two ground glass screens (21 × 16 cm.) in two well-lit "corridors" formed by walls of heavily textured cloth (Figs. 2 and 3). The screens and projectors were mounted on trolleys which could be pulled along rails, thus varying the distances of the stimuli from the animal. The sizes of the discs were controlled by iris diaphragms in the projectors. The animal sat in a transport cage in front of two Perspex windows (20 × 12½ cm.) looking on to the two corridors.

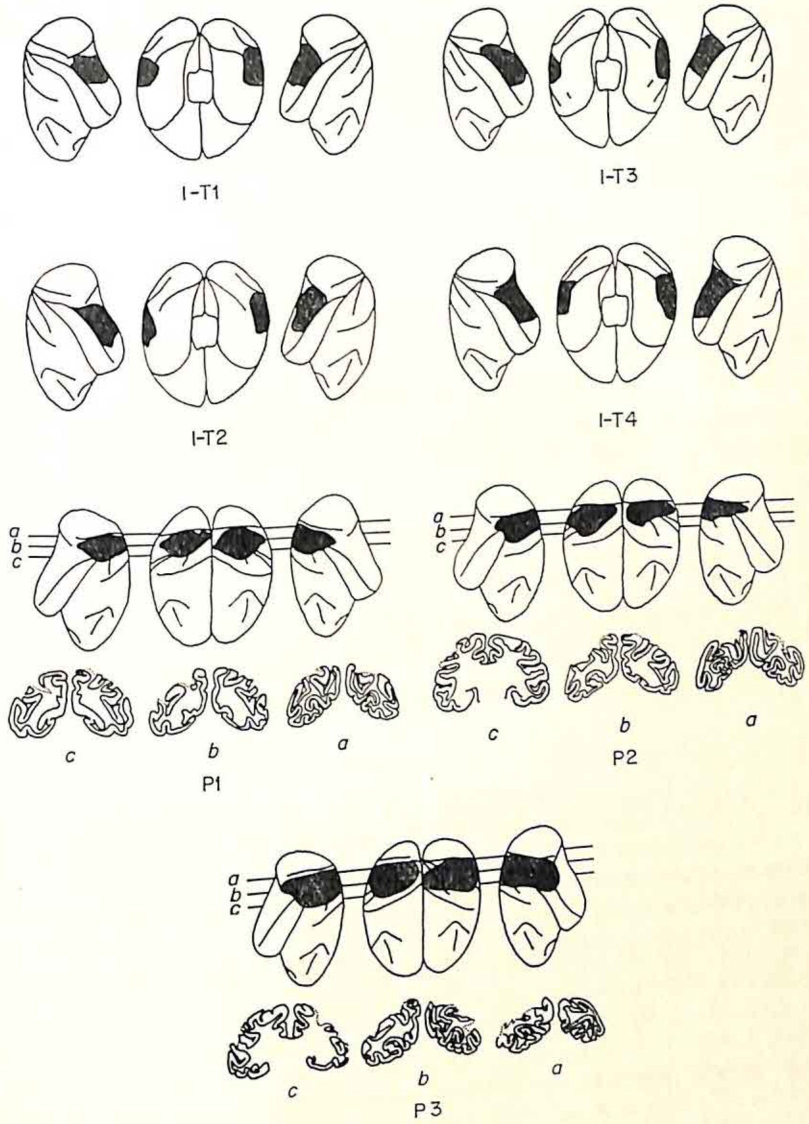


FIGURE 1.

Except when the projection screens were closest to him the animal was unable to see them both at once so that to compare the stimuli he had to look down one corridor after the other. He made his choice of the larger disc by pressing the window in front of the appropriate corridor. A correct choice was rewarded by a peanut delivered automatically to a trough just below the window; an incorrect choice was mildly punished by the room lights being switched off for a few seconds. In between trials a wooden screen covered the windows. At the beginning of a trial the screen was raised allowing the animal to see the stimuli. When he had made his choice the screen was lowered—after a short delay in the case of a correct choice, immediately in the case of an incorrect one. The sizes of stimuli used were 4, 2 and 1-cm. diameter. They were paired either as 4 cm. against 2 cm. or 2 cm. against 1 cm., so that the larger disc was always twice the diameter of the smaller. The distances used covered the range from 30 to 120 cm. from the animal (it being assumed for these measurements that he sat with his head almost up against the bars of the cage).

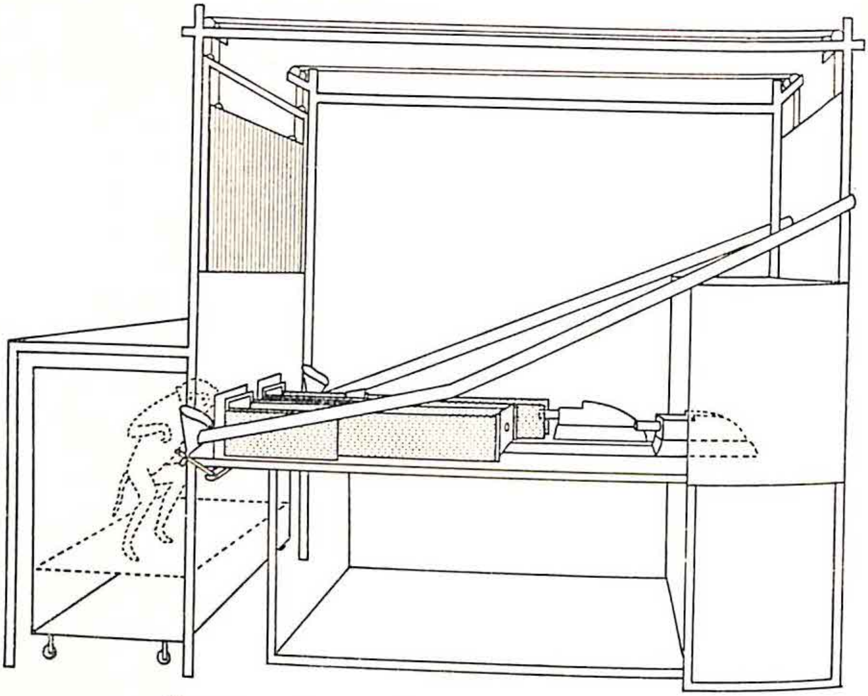


FIGURE 2. General view of the testing apparatus.

Pre-operative testing schedule

Thirty trials, with correction, were given per day on 5 or 6 days a week. The correct window (the window in front of the larger disc) and the stimulus pair used (4 cm. vs. 2 cm. or 2 cm. vs. 1 cm.) were determined by random schedules. The discrimination proved to be difficult to learn and a long period of shaping was required. Initially (pre-op., Stage I), the stimuli were presented at the near ends of the corridors, up against the windows, and the animals were trained to a criterion of 27/30 correct on 1 day. The stimuli were then presented at *equal* distances increasingly further down the corridors and the animals trained again to a criterion of 27/30 correct whatever the distance. Trials in which the stimuli were at *unequal* distances (pre-op., Stage II) were then introduced according to the following schedule (see Table I). Each block of 30 trials contained 18 "equal distance trials" and 12 "unequal distance trials." The unequal distance trials were presented alternately with

TABLE I

Schedule of trials within a single 30-trial block

Class of trial	Positions of stimuli (cm.)		Number per block
	large	small	
Equal	30–120	30–120	18
L4X	120	30	2
L2X	120	60	2
	60	30	2
S4X	30	120	2
S2X	60	120	2
	30	60	2

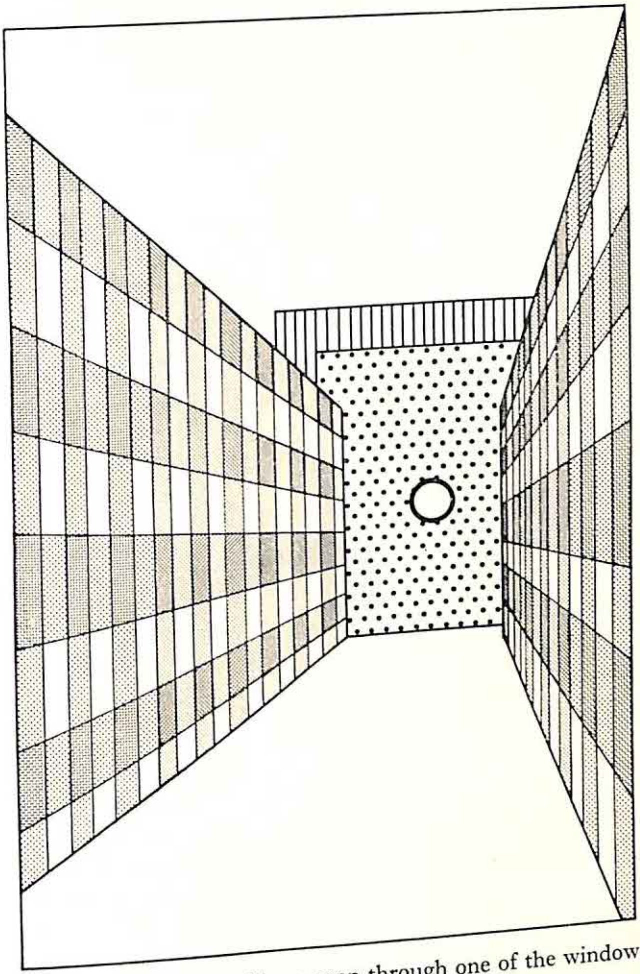


FIGURE 3. Stimulus disc as seen through one of the windows.

equal distance trials in the middle of the block. The distances of the stimuli on the equal distance trials were varied randomly over the whole range from 30 to 120 cm. The distances on the unequal distance trials were chosen as shown in the table so as to give 4 trials in which the larger stimulus was twice the distance of the smaller (L2X), 2 trials in which the larger stimulus was 4 times the distance of the smaller (L4X), 4 trials in which the smaller stimulus was twice the distance of the larger (S2X), and 2 trials in which the smaller stimulus was 4 times the distance of the larger (S4X), the order of these trials being random. The animals were trained on this schedule to a criterion of better than 85 per cent over 7 days on each of the three main classes of trial, "equal distance," "large far" (L4X and L2X) and "small far" (S4X and S2X), taken separately.

Post-operative testing schedule

Testing was resumed on the seventh day after operation in all animals except P3, who did not recover sufficiently until the eleventh day. Initially (post-op., Stage I), the operated animals were given only equal distance trials, distance varying over the whole range. All except IT4 were retrained to a criterion of 27/30 on these before going on to unequal distance trials. For reasons which will become clear later (the necessary failure on equal distance trials of an animal adopting the "distance strategy") it was decided to start IT4 on unequal distance trials after only 3 days although he had not reached the criterion. The unequal distance trials (post-op., Stage II) were introduced according to the same schedule as before operation. For the first 7 days of testing with unequal distance trials both correct and incorrect choices on these trials were rewarded; thereafter, selective reinforcement was reintroduced and maintained. Testing was continued until the animal reattained the pre-operative criterion of 85 per cent over 7 days on the three main classes of trial, or until 35 days testing were completed.

Results

Pre-operative discrimination learning

All animals took a very long time to reach the final criterion, total days to criterion ranging from 20 for IT1 to 162 for IT4 (see Table II). The initial discrimination with the stimuli up against the windows was learned in a few hundred trials, but when the stimuli were moved back along the corridors performance dropped almost to a chance level. This stage, involving only equal distance trials, was eventually mastered after considerably more training, but when the final stage involving unequal distance trials began performance relapsed in all animals except IT1 and P1. The relapse occurred on both equal and unequal distance trials although the animals had previously performed well on the former. With further training performance improved, more or less in parallel on both kinds of trial, but in several animals it remained just short of criterion for many days: IT4, for instance, scored over 75 per cent on all three major classes of trial for over 35 days before reaching 85 per cent.

TABLE II

Total days to criterion (excluding the criterion run) at different pre- and post-operative stages

Subject	Pre-op.		Post-op.	
	Stage I	Stage II	Stage I	Stage II
P1	46	8	2	3
P2	45	26	1	1
P3	89	30	10	0
IT1	20	0	12	7
IT2	67	21	2	> 35
IT3	85	39	2	> 35
IT4	48	114	> 3	> 35

Post-operative general behaviour

(i) *Parietals.* The three posterior parietal animals were obviously severely impaired in the first hours following operation. They stumbled around their cages, bumping into the walls and hitting their heads on protruding objects; if they tried to steady themselves by catching hold of the wire, they reached too short, too far or in the wrong direction, missed their target and fell; they appeared quite insensitive to hard pinches on one or both sides of their bodies and often let their limbs droop lifelessly; they made no attempt to reach for food and showed no interest in it if it was put into their hands although they took it greedily if it was pressed against their lips. At the same time, automatic movements—grooming, etc.—were nearly normal, and visual recognition, judged by reactions to familiar objects and people, was unimpaired. Recovery from these gross behavioural disorders was rapid. P1 and P2 appeared more or less normal after 4 days and P3 after 8 days, although all three remained slow and clumsy and showed severe misreaching for some days more.

(ii) *Inferotemporals*. None of the inferotemporal animals showed any obvious signs of disturbance after the operation.

Post-operative discrimination performance

(i) *Parietals*. Despite a residual motor disturbance which handicapped them in picking nuts out of the troughs, P1 and P2 regained criterion on equal distance trials almost at once, taking 1 and 2 days, respectively. P3 took 10 days but this apparent impairment was due at least in part to the much greater difficulty this animal had in picking out nuts, which for the first 2 days made him leave more than half of them behind and so upset the reinforcement schedule. When the unequal distance trials were introduced the animals showed almost no deficit and the pre-operative criterion was achieved in 3, 1 and 0 days, respectively (see Table II).

(ii) *Inferotemporals*. IT1 performed very erratically at first on equal distance trials and for 10 days refused to complete 30 trials, although he regained criterion eventually in 12 days. IT2 and IT3 regained criterion at this stage quickly, taking 2 days each. IT4 performed at near chance level for 3 days and, as mentioned earlier, was not taken to criterion. When the unequal distance trials were introduced all four animals performed very badly. IT1 achieved criterion only after 7 days and the other three failed to do so in 35 days. Figure 4 shows the inferotemporals' performance at Stage II in terms of scores over successive 7-day periods on each of the three main classes of trial. It will be seen that there was not a general breakdown, performance on different classes of trial being affected to quite different extents. The pattern of deficit varied from animal to animal. IT1

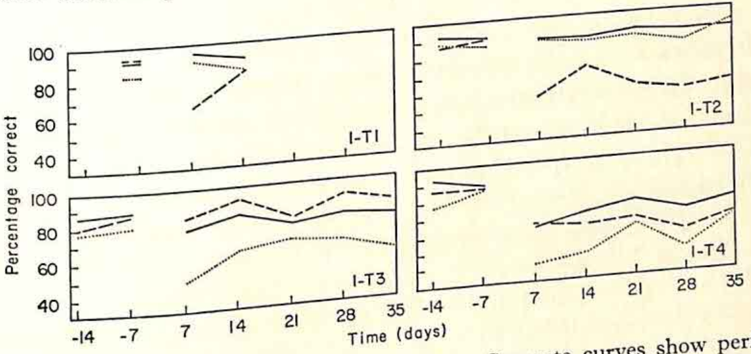


FIGURE 4. Performance curves for IT1, 2, 3 and 4. Separate curves show performance on "equal distance," "large far" (L2X and L4X) and "small far" (S2X and S4X) trials. Days −14 and −7 correspond to the period just before operation and days +7 through +35 to the period of post-operative testing at Stage II. —, Equal distance; - - -, large far; · · ·, small far.

showed a drop in his score on the large far trials but relatively little change on equal distance and small far trials. IT2 showed a similar drop on large far trials with little change on equal distance and small far trials. IT3 showed a big drop on small far trials, a smaller drop on equal distance trials and if anything an improvement on large far trials. IT4 showed a big drop on small far trials and smaller drops on both equal distance and large far trials (it is noteworthy that this animal who had never scored under 75 per cent in the 35 days immediately preceding operation never scored over 75 per cent post-operatively). The performance of IT1 recovered quite quickly but IT2, IT3 and IT4 showed a remarkably stable impairment. When the scores on the unequal distance trials are broken down into total post-operative scores on the classes L4X, L2X, S4X and S2X (see Table III) the differences between the animals and the heterogeneity of their scores are more marked still.

TABLE III

Average scores (per cent correct) throughout post-operative Stage II

Subject	Equal	L4X	L2X	S4X	S2X
IT1	94	69	79	81	91
IT2	90	46	67	87	88·5
IT3	82·5	85·5	92	61·5	65
IT4	70	49	69	51	51

Analysis of the inferotemporals' performance

The post-operative performance of the inferotemporals, each animal showing a particular rather regular distribution of scores, could not easily be attributed to a non-specific disability. The evidence suggests—and this impression was supported by observation of the animals during the testing sessions—that even in making errors they continued to respond discriminatively to the stimuli, but that the basis for their discrimination had changed. This being so, then (i) Why should the patterns of performance have been just such as they were?, (ii) Why, in three animals at least, should these patterns have remained so stable over 35 days' testing, apparently uninfluenced by differential reinforcement?

In line with the theoretical aim of this study, the possibility was examined of a defect in perceptual size constancy. The hypothesis was that post-operatively the inferotemporals continued to respond to the perceived size of the stimuli, choosing the stimulus which looked the larger, but that now perceived size no longer corresponded to physical size because of some error in the computation of size from the information at the retina. This hypothesis was made the basis for a specific analysis.

The reasoning went as follows. Veridical perception of size depends on combining information about both the retinal image size of an object and its distance according to the rule

$$\text{perceived size} = k\,(\text{image size} \times \text{distance}).$$

If the mechanism which normally applies this rule were to be put out of action, a form of compromise might be adopted, viz. perceived size might be made dependent not on both image size and distance but on one alone with the other arbitrarily fixed. The two possible strategies would be (i) to use information about image size and assume an arbitrary value for distance, giving

$$\text{perceived size} = \text{image size} \times \text{constant}$$

or (ii) to use information about distance and assume an arbitrary value for image size, giving

$$\text{perceived size} = \text{constant} \times \text{distance}.$$

In an animal whose constancy mechanism were to be functioning badly but not permanently out of order, perception might fluctuate between the use of the proper rule and these two partial strategies. Suppose the "image size strategy" were to be adopted a proportion p of the time, the "distance strategy" a proportion q of the time and the "correct strategy," combining both size and distance information, a proportion $1\text{-}p\text{-}q$ of the time, then it should be possible to make precise predictions about performance on a size discrimination involving known sizes and distances. The predicted scores for the different classes of trial in the present experiment would be as follows.

Equal distance. Here the image size of the larger stimulus is twice that of the smaller and the distances are equal. Hence with the image size strategy the animal would score 100 per cent, with the distance strategy he would score at chance, i.e. 50 per cent, and with the correct strategy he would score 100 per cent. His average score would thus be

$$\begin{aligned}\text{per cent correct} &= 100p + 50q + 100\,(1\text{-}p\text{-}q)\\ &= 100 - 50q.\end{aligned}$$

L4X. Here the image size of the larger stimulus is half that of the smaller and the distance of the larger is 4 times that of the smaller. Hence with the image size strategy the animal would score 0 per cent, with the distance strategy he would score 100 per cent and with the correct strategy he would score 100 per cent. His average score would thus be

$$\begin{aligned}\text{per cent correct} &= 0p + 100q + 100\,(1\text{-}p\text{-}q)\\ &= 100 - 100p.\end{aligned}$$

L2X. Here the image size of the larger stimulus is equal to that of the smaller and the distance of the larger is twice that of the smaller. Hence with the image size strategy the animal would score 50 per cent, with the distance strategy he would score 100 per cent and with the correct strategy he would score 100 per cent. His average score would thus be

$$\begin{aligned}\text{per cent correct} &= 50p + 100q + 100\,(1\text{-}p\text{-}q)\\ &= 100 - 50p.\end{aligned}$$

S4X. Here the image size of the larger is 8 times that of the smaller and the distance of the larger is a quarter that of the smaller. Hence with the image size strategy the animal would score 100 per cent, with the distance strategy he would score 0 per cent and with the correct strategy he would score 100 per cent. His average score would thus be

$$\text{per cent correct} = 100p + 0q + 100\,(1\text{-}p\text{-}q)$$
$$= 100 - 100q.$$

S2X. Here the image size of the larger is 4 times that of the smaller and the distance of the larger is a half that of the smaller. Hence with the image size strategy the animal would score 100 per cent, with the distance strategy he would score 0 per cent and with the correct strategy he would score 100 per cent. His average score would thus be

$$\text{per cent correct} = 100p + 0q + 100\,(1\text{-}p\text{-}q)$$
$$= 100 - 100q.$$

In summary, the predicted average scores would be

Equal	$100 - 50q$
L4X	$100 - 100p$
L2X	$100 - 50p$
S4X	$100 - 100q$
S2X	$100 - 100q$.

With this hypothetical analysis in mind, an attempt was made to find values of p and q which, substituted in the equations, give expected scores close to the observed average scores over the post-operative period. This meant trying to fit the scores on *five* different classes of trial by assigning values to *two* independent variables (with the constraint that $p + q \leqslant 1$): a good fit should be obtainable only if the animal's performance did conform to the predicted pattern.

TABLE IV

Observed scores (O) for post-operative Stage II compared with expected scores (E) calculated for the values of p *and* q *shown*

Subject	p	q	Equal O	Equal E	L4X O	L4X E	L2X O	L2X E	S4X O	S4X E	S2X O	S2X E
IT1	0·36	0·14	94	93	69	64	79	82	81	86	91	86
IT2	0·58	0·15	90	92·5	46	42	67	71	87	85	88·5	85
IT3	0·14	0·36	82·5	82	85·5	86	92	93	61·5	64	65	64
IT4	0·50	0·50	70	75	49	50	69	75	51	50	51	50

The result of the attempt is shown in Table IV; the same is shown diagrammatically in Figure 5. It will be seen that the rank order of the expected scores is right in each case, while the mean absolute discrepancy between expected and observed scores is under 3 per cent. The closeness of this fit clearly gives strong support to the hypothesis.

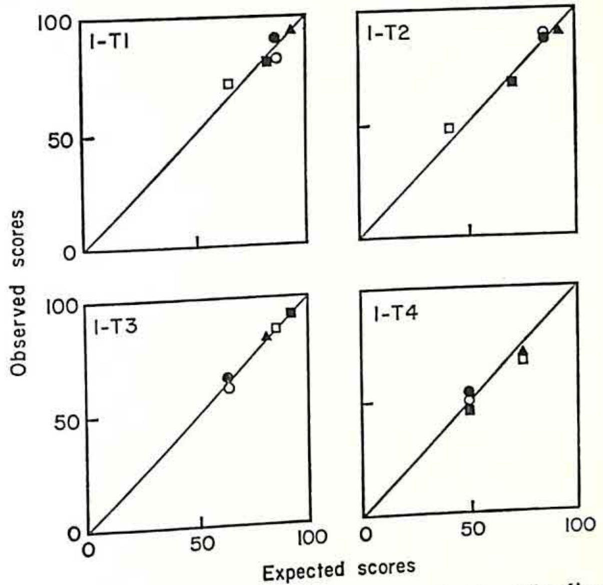


FIGURE 5. The same data as in Table IV, shown diagrammatically. The diagonals are the "lines of perfect fit." ▲, Equal; ■, L4X; □, L2X; ●, S4X; ○, S2X.

Analysis of the pre-operative performance

Although the analysis was developed to explain the performance of the inferotemporals, its success prompted an attempt to treat the pre-operative data in the same way. As already described, several animals before operation took a long time to master the stage of the problem involving unequal distance trials, despite having reached criterion with equal distance trials alone. It is at least a possibility that their poor performance was due to a natural difficulty with the discrimination, similar to the inferotemporal deficit but much less severe. This was tested by trying to find values of p and q which fit with the pre-operative scores. Only the final 14 days up to and including those to the final criterion were considered, since in several animals earlier performance appeared very erratic, and in any case some required little more than 14 days altogether. Taking each animal individually, the best fit obtainable is in several cases very poor. This could be due, however, to the relatively small number of trials involved in 14 days testing, being insufficient for the statistical norms to emerge. If the data from all 7 animals are pooled to give more weight to the scores (as is permissible since the equations are linear), the fit is unexpectedly good (Table V). Further consideration was therefore given to the

TABLE V

Observed scores (O) pooled for all animals over the final 14 days before operation compared with expected scores (E) calculated for the values of p *and* q *shown*

		Equal		L4X		L2X		S4X		S2X	
p	q	O	E	O	E	O	E	O	E	O	E
0·24	0·17	89·5	91·5	78·5	76	86·5	88	82	83	85	83

one animal, IT4 ,whose performance by contrast with that of the others was fairly stable for much longer before he reached criterion. When IT4's scores over the 35 days up to and including those to criterion are considered, the fit turns out to be very good (Table VI). It seems therefore that at least the terminal performance of the un-operated animals followed the same pattern as the performance of the inferotemporals.

TABLE VI

Observed scores (O) for IT4 over the final 35 days before operation compared with expected scores (E) calculated for the values of p *and* q *shown*

		Equal		L4X		L2X		S4X		S2X	
p	*q*	O	E	O	E	O	E	O	E	O	E
0·26	0·23	87·5	88·5	77	74	85	87	78·5	77	79	77

Discussion

The evidence of the present study may be summarized as follows: (1) Normal monkeys could be trained to choose the larger of two projected stimuli independently of distance, although they took a long time to reach a consistently high level of performance. (2) Large posterior parietal lesions, after pre-operative training, caused no deficit on this task. (3) Inferotemporal lesions, by contrast, caused a severe breakdown, from which three out of four animals showed almost no recovery. (4) The pattern of performance after inferotemporal lesions can be described as due to a tendency to oscillate between a "correct" and two "incorrect strategies." (5) This tendency was also apparent to at least some extent in the unoperated animals.

The analysis of the deficit in terms of correct and incorrect strategies rose directly out of the hypothesis of a breakdown in perceptual size constancy. On this hypothesis the strategies represent alternative modes of perceptual organization, size perception being made to depend either on both retinal size and distance information or on retinal size or distance information alone. A problem is raised, however, by the finding that the analysis applies not only to the performance of the inferotemporals but to the performance of the unoperated animals as well. Although the pre-operative "deficit" was relatively slight, the fact that it followed the same pattern as the post-operative deficit suggests that it had the same cause and that therefore it too should be attributed to a failure of perceptual constancy. This suggestion, without external support, is hard to take. It is certainly unlikely that healthy free-ranging monkeys can be subject under natural conditions to sporadic errors in size perception. Two considerations, however, make the possibility of a perceptual disorder in the unoperated monkeys of the present study not implausible. First, these monkeys were manifestly not normal free-ranging animals, but captive animals confined to small laboratory cages where their scope for "active" visual exploration, such as is required for the maintenance of adequate perception was severely restricted (cf. Held and Hein, 1964). Second, the situation in which the animals were tested, looking through windows at meaningless projected stimuli at the ends of corridors giving rather limited depth cues, was highly artificial and such as might emphasize any potential perceptual instability.

The necessity for special pleading to account for the performance of the "normal" monkeys must make the hypothesis that the inferotemporals suffered a specific impairment of perceptual constancy rather less attractive. It is, however, hard to see how the results can be otherwise explained. There have certainly been other theories of the inferotemporal defect which, resting on earlier evidence for a disturbance in perception, have proposed explanations in terms of visual field defects, or impairment of selective attention or perceptual categorization (for a recent summary see Wilson, 1968). The results of the present experiment do not at all easily fit into the framework of these previous theories. It seems possible, rather, that the hypothesis of impaired constancy may itself be able to cover many of the earlier findings, while having the advantage of accounting in detail for the disturbance of size perception found in this study. This hypothesis has moreover the advantage of tying in the inferotemporals' deficit with an established clinical syndrome in man, for abnormalities of size perception occur quite commonly among the symptoms of perceptual disturbance referred to already. The character of the clinical disorder can be illustrated from case records, as, for instance, in a case of diffuse cortical damage after head injury where "On maintained fixation the objects, after a latent period of 5 to 6 sec., apparently grew rapidly to more than double their original size" (Willanger and Klee, 1966). Similar phenomena, in other pathological conditions, are reported by Penfield and Jasper (1954), Chapman (1966) and Heron *et al.* (1956), among others. In the absence of proper experimental study of such cases, parallels between man and monkey must remain conjectural. But on the present evidence the hypothesis of a breakdown in constancy after inferotemporal lesions in monkeys, with a parallel in such human cases may be considered at least in intriguing possibility.

Still to consider is the failure to find any specific impairment in the control animals of this study, with parietal lesions. In view of the manifest spatial disorientation of these animals in the first few days after operation, it is somewhat surprising that they should have shown no deficit at all in size constancy. It should be noted however that the most obvious signs of disorientation had disappeared by the time formal testing was started, so that a transient perceptual disturbance may have been missed. It is perhaps wisest simply to record our negative finding, while withholding judgement on the more general question of what, if any, permanent deficit was present in these animals.

It is a pleasure to acknowledge the help of Dr. A. Cowey, who performed the surgery on four of the monkeys, and also the technical assistance of Mr. R. Hutchison and Miss S. Hopkins. Professor O. L. Zangwill made research facilities available in Cambridge and gave valuable encouragement. The work was supported by an apparatus grant from the Medical Research Council (965/96/B) and a Research Scholarship to the first author from the Science Research Council.

References

Chapman, J. (1966). The early symptoms of schizophrenia. *Br. J. Psychiat.* **112**, 225–51.

Critchley, M. (1953). *The Parietal Lobes.* London: Edward Arnold.

Hecaen, H., Penfield, W., Bertrand, C. and Malmo, R. (1956). The syndrome of apractagnosia due to the lesion of the minor cerebral hemisphere. *Archs. Neurosurg. Psychiat.* **75**, 400–34.

Held, R. and Hein, A. (1963). Movement-produced stimulation in the development of visually-guided behavior. *J. comp. physiol. Psychol.* **56**, 872–6.

Heron, W., Doane, B. K. and Scott, T. H. (1956). Visual disturbances after prolonged perceptual isolation. *Can. J. Psychol.* **10**, 13–8.

Holway, A. H. and Boring, E. G. (1941). Determinants of apparent visual size with distance variant. *Am. J. Psychol.* **54**, 21–37.

Mayer-Gross, W., Slater, B. and Roth, M. (1960). *Clinical Psychiatry.* London: Cassell & Co.

Penfield, W. and Jasper, H. (1954). *Epilepsy and the Functional Anatomy of the Human Brain.* Boston: Little, Brown & Co.

Willanger, R. and Klee, A. (1966). Metamorphopsia and other visual disturbances with latency occurring in patients with diffuse cerebral lesions. *Acta neurol. Scand.* **42**, 1–18.

Wilson, M. (1968). Inferotemporal cortex and the processing of visual information in monkeys. *Neuropsychologia* **6**, 135–40.

Received 4 March 1969

Q. Jl exp. Psychol. (1969) **21**, 239–247

THE EFFECT OF OBJECT FAMILIARITY ON THE PERCEPTION OF SIZE AND DISTANCE

WALTER C. GOGEL
University of California, Santa Barbara

The perceived sizes and perceived distances of familiar objects were investigated in two experiments in which images of familiar objects were presented monocularly, one at a time, in an otherwise dark field of view. It was found that the angular size of the objects as well as their familiar size determined reported size. Reported distances were increasingly underestimated as a function of increasing simulated distances of the objects. The results are consistent with the conclusion that, as a function of the retinal size of the objects, the observer perceives the familiar objects as off-sized, and, that as a consequence of these off-sized perceptions, the observer's judgements of the object distances reflect inferential rather than perceptual processes.

Introduction

A previous study (Gogel and Newton, 1969) has indicated that the perceived size of a familiar object presented monocularly in an otherwise dark visual field is as much determined by visual angle as by familiar size. The explanation for this result is in terms of a tendency called the specific distance tendency (Gogel, 1969*b*). This tendency is supported by experimental data indicating that an object with no cues to size or distance will be perceived at a specific distance (usually 4 to 12 ft. from the observer), and will acquire a perceived size at that distance depending upon its angular size. It can be concluded that two sources of information as to size are available, when a familiar object is presented under otherwise reduced visual conditions. One source of perceived size occurs from retinal size by means of the specific distance tendency. The second source is the familiar size of the object. When the perceived sizes resulting from these sources are disparate, the observer perceives an off-sized object rather than a normal-sized object. The question can be raised as to the distance judgments that occur when the observer is confronted with a perceptually off-sized object. Research has suggested that the observer expects an object to appear smaller when it is distant rather than near (Jenkin and Hyman, 1959; Carlson, 1960, 1962*a*, *b*; Epstein, 1963). It is likely, therefore, that the observer will report that the perceptually off-sized object is distant or near, depending upon whether the object is perceived as smaller or as larger than normal. The distance reports made under these conditions probably should be considered to be inferential rather than perceptual, i.e. more determined by the tendency to report that an object is distant because its perceived size is small than to report the distance actually perceived. The purpose of the present study is (1) to verify the influence of retinal size on the perceived size of familiar objects and (2) to investigate the factors determining the "inferential" judgements of distance resulting from the perceptions of size.

Experiment 1

Apparatus

The apparatus was the same as that used in the study by Gogel and Newton (1969). Coloured positive transparencies of familiar objects were presented one at a time to the right eye at a distance of 133 cm. from the observer. The transparencies were mounted on a light box and were transilluminated by diffused white light from a fluorescent source. The observer viewed the transparencies monocularly by means of a first surface mirror through a + 0·75 spherical lens placed immediately in front of his right eye. The lens removed the possible accommodative cue to the physical distance of the transparencies by optically locating the familiar objects at an infinite distance. The familiar objects appeared directionally in front of the observer in an otherwise totally dark visual field. The dark observation position contained a head and chin rest and a shutter that was raised to present the stimuli. At no time during the experiment was anything visible except the transparencies of the familiar objects. The stimuli presented were identical visually to real objects presented at the simulated distances, but with the accommodation for a far distance. No cues of any kind were available to the observer to indicate that transparencies rather than the real objects were being observed, or that the physical distance of the transparencies was 133 cm. Communication between the experimenter and observer occurred by means of a headphone and a microphone with a continuous masking noise presented to the observer through the headphones at all other times.

The transparencies were the same as those used previously (Gogel and Newton, 1969). The names of the familiar objects, their widths S_T on the transparencies, the physical width S_F of the familiar object photographed, the visual angle θ of the width of the transparency image, and the distance D_F simulated by each transparency are given in the first five columns of Table I. The simulated distance D_F is the physical distance from the observer at which a normal-sized object of that familiar kind would have to be placed in order to subtend the same visual angle as the transparency.

Observers

The observers were 32 students in an introductory course in psychology. All had a visual acuity of at least 20/20 in their right eye, as measured with the Landolt-Ring test of the Keystone Orthoscope and were naive with regard to the purpose of the experiment.

Procedure

The observer was instructed to indicate the perceived widths of the objects and the perceived distances of the objects from his eyes by verbal reports expressed in feet or inches or in a combination of both. At no time during the experiment was the observer given any indication that transparencies were being used to simulate the objects. Since some of the objects were irregular in shape the observer was asked to imagine a vertical line on the left most part of the object and another on the right most part of the object in making his report of apparent width. The width to be indicated was the apparent horizontal distance between these two imaginary lines. Before being presented with any of the familiar objects an observer was dark adapted for 3 min. With half of the observers the apparent size responses preceded the apparent distance responses and for the remaining half the order was reversed. Both a size and distance report were obtained from one familiar object before presenting another. The order of presenting the familiar objects differed for each observer as determined by a balanced latin-square design.

Results

The mean and median results obtained from the width (S') and distance (D') reports converted to centimetres are given in four of the columns of Table I.

TABLE I

Characteristics of the familiar objects and results obtained from the size and distance reports

Object names	Width of object on transparency S_T (cm.)	Width of familiar object S_F (cm.)	Visual angle of image width on transparency θ (rad.)	Simulated distance D_F (cm.)	Mean reported width S' (cm.)	Median reported width S' (cm.)	Computed off-sized perceptions from means S'/S_F	Computed off-sized perceptions from medians S'/S_F	Mean reported distance D' (cm.)	Median reported distance D' (cm.)
Guitar	5·30	96·5	0·0398	2425	30·8	7·0	0·32	0·07	853	610
Saw	9·80	74·9	0·0737	1017	29·5	9·5	0·39	0·13	436	366
Tennis racquet	10·50	68·9	0·0789	873	29·6	11·1	0·43	0·16	495	411
Magazine	4·00	21·6	0·0301	683	9·9	3·8	0·46	0·18	413	381
Cigarette carton	5·70	28·6	0·0429	666	14·6	7·6	0·51	0·27	379	305
Telephone	6·50	21·0	0·0489	428	14·6	11·4	0·70	0·54	260	244
Ruler	11·35	30·5	0·0853	357	15·5	10·2	0·51	0·33	232	206
Candy Bar	5·60	12·1	0·0420	287	8·8	7·6	0·73	0·63	158	130
Sunglasses	9·60	13·4	0·0722	185	10·9	12·1	0·81	0·90	110	91
Playing card	5·00	5·7	0·0376	151	6·4	5·1	1·12	0·89	106	91
Cough drop box	11·40	10·0	0·0857	117	20·3	10·2	2·03	1·02	82	53
Tape dispenser	9·20	7·4	0·0692	107	8·3	7·6	1·12	1·03	82	61
Stamp	3·30	2·2	0·0240	88	3·0	2·5	1·36	1·14	97	61
Razor blade cont.	11·40	6·6	0·0857	77	9·1	6·4	1·38	0·97	59	46
Key	11·35	5·4	0·0853	63	9·1	5·1	1·69	0·94	50	30
Penny	4·60	1·9	0·0350	54	3·8	1·9	2·00	1·00	52	30

Pearson product-moment correlation coefficients were computed between values of perceived size S' and either visual angle θ or familiar size S_F for both mean and median values of S'. The resulting values of r were as follows: for means, $r_{S', \theta} = 0{\cdot}34$, $r_{S', S_F} = 0{\cdot}92$; for medians, $r_{S', \theta} = 0{\cdot}59$, $r_{S', S_F} = 0{\cdot}36$. Only the 0·92 and 0·59 values were significant at the 0·02 level ($t_{14} = 1{\cdot}36$, $8{\cdot}78$, $2{\cdot}74$ and $1{\cdot}44$, respectively). Median values of S' were more highly correlated with θ, while mean values of S' were more highly correlated with S_F. The correlation between the physical values S_F and θ was $r = 0{\cdot}07$. The differences between the means and medians in Table I suggests that, although for some observers perceived width was mainly determined by retinal size, for others perceived width was mainly determined by familiar size. It should be noted in Table I that the mean and median values of perceived width for the objects of large familiar width (large S_F) were severely underestimated relative to their familiar width. On the other hand, for the objects of small familiar width (small S_F) mean but not median reported values tended to be large relative to familiar width. This tendency to underestimate objects of large familiar size and to a lesser extent to overestimate objects of small familiar size indicates that the objects of large familiar size tended to be perceived as small-sized versions of familiar objects and the objects of small familiar size tended to be perceived as large-sized versions of familiar objects. This is reflected in the entries in the two columns labelled "Computed Off-Sized Perceptions" which were derived by dividing the experimentally obtained (mean or median) values of perceived width S' by the width S_F of the familiar object. The values throughout these two columns would have been unity if familiar size alone were the sole determiner of perceived size. It is obvious that retinal size as well as familiar size was an important determiner of perceived size in this experiment.

The results from the distance reports are given in the last two columns of Table I. Pearson product-moment correlations between simulated distances D_F and either mean or median reported distance D' ($r = 0{\cdot}97$ and $0{\cdot}84$) were significant at the 0·01 level ($t = 14{\cdot}82$ and $5{\cdot}84$, respectively). It will be noted that, although D' clearly tended to increase with an increase in D_F, the value of reported distance was increasingly underestimated with increasing values of D_F. This is demonstrated by the slope of less than 45° for the lines of best fit applied to the relation between D_F and either mean or median reported distance. These equations are:

$$D'_{\text{mean}} = 0{\cdot}36 D_F + 72{\cdot}46 \text{ cm.}$$
$$D'_{\text{mdn.}} = 0{\cdot}24 D_F + 63{\cdot}56 \text{ cm.}$$

Although it is clear that the underestimation of the simulated distance increased with the increase in the magnitude of the simulated distance, this result must be interpreted with caution. It is possible that the underestimation demonstrated here, rather than being related to the judgements of distance associated with perceptually off-sized objects, occurs whenever numbers are applied in the estimation of distance. In support of this possibility, it has consistently been found that objects presented under full cue conditions of observation, i.e. under conditions in which large errors in reported distance would not be expected to occur, also tend to be increasingly underestimated with increasing distance (Gogel, 1968*a;* Gogel and Mertens, 1966; Gogel, Hartman, and Harker, 1957). Verbal underestimations of distance, therefore, possibly can be attributed to processes involved in the application of numbers to distances and do not necessarily reflect judgemental or perceptual underestimations. The question occurs whether the underestimations found in the present study can be explained by such a methodological artifact. The purpose of Experiment 2 was to investigate this problem.

Experiment 2

Apparatus

The apparatus for presenting the transparencies in Experiment 2 was identical to that used in Experiment 1. But only 10 of the 16 transparencies of Experiment 1 were used in Experiment 2. These, presented one at a time in an otherwise totally dark visual field, were the guitar, tennis racquet, saw, cigarette carton, telephone, ruler, candy bar, sunglasses, tape dispenser and key. In addition to the apparatus for presenting the transparencies, a full cue situation also was used. This consisted of an alley (approximately 3 ft. wide by 29 ft. long) with a white and a black stripe (each 18 in. wide) running the length of the alley. The walls of the alley were formed by white cloth. Six familiar objects (a box of cough-drops, a bottle of ink, a cup, a soup can, a pencil sharpener and a flashlight) were presented in the alley at 4, 7, 11, 14, 21 and 28 ft., respectively, from the observer. The entire alley was illuminated by a series of overhead lights and was viewed monocularly. When the alley was presented, the lens in the viewing position was removed and the observer was asked to report the distance from his eye to each of the six familiar objects in the alley. The transparencies were never visible during the presentation of the alley. Only distance, not size, reports were obtained in Experiment 2.

Observers

The observers were 32 students enrolled in a class in introductory psychology. A visual acuity of at least 20/20 in their right eye, as measured with the Landolt-Ring test of the Keystone Orthoscope was required. None of these observers had participated in Experiment 1 and all were naive with regard to the purpose of the experiment.

Procedure

The series of ten transparencies always was presented to the observer before presenting the full-cue alley. As in Experiment 1, the observers were dark adapted for 3 min. before being presented with the first transparency. The order of presenting the different transparencies was different for each observer as determined by a balanced latin-square design. The order of judging the apparent distance to the objects in the full-cue alley was randomly determined with a different random order for each observer.

Results

The average reports D' of distance from Experiment 2 are shown in Figure 1 as a function of the simulated and the physical distance D of the objects. In the case of the transparencies, D is the simulated distance of the objects and in the case of the alley, D is the physical distance of the objects in the alley. The dashed line in Figure 1 represents the results that would have been obtained if the simulated or physical distances had been correctly reported. In Figure 1 both of the curves using the obtained data are below the dashed line with the curve from the average values of D' using the transparencies (reduced condition curve) lying below the corresponding curve obtained from the alley (full-cue condition curve). Straight lines of best fit were calculated by the method of least squares from the verbal

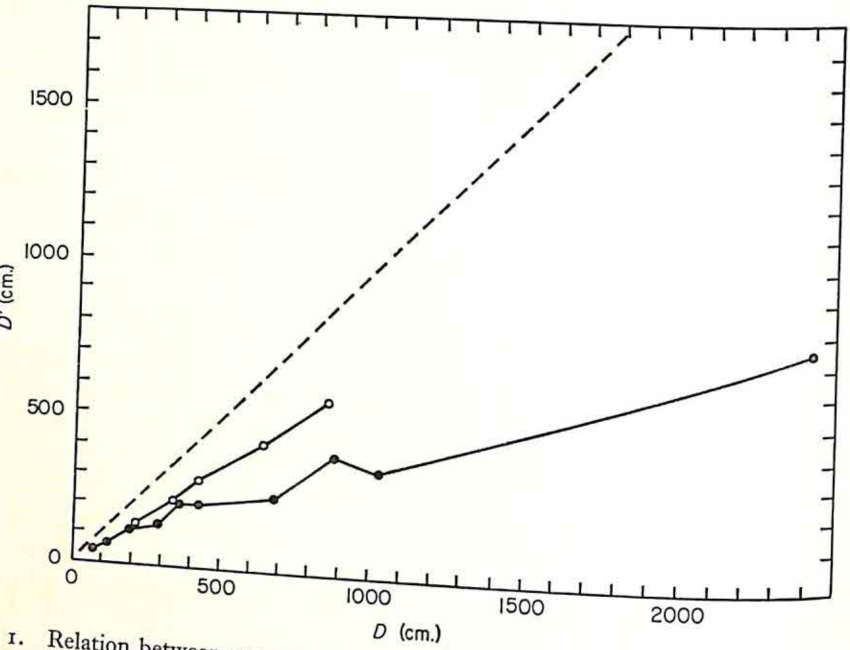


FIGURE 1. Relation between average reported distance D' and simulated (or physical) distance D from the full-cue (alley) condition (○) and also from the reduced (transparency presentations) condition (●). The dashed line represents the results expected if reported distance were equal to physical distance using the alley or to simulated distances using the transparencies.

reports of distance from each observer both for the transparency presentations and for the full-cue alley. Averaging the slopes and intercepts of these lines of best fit gives $D' = 0{\cdot}69D - 12{\cdot}0$ cm. for the alley and $D' = 0{\cdot}31D + 53{\cdot}9$ cm. for the transparency presentations. The slope of the latter equation is significantly less than that of the former at the 0·01 level ($t = 5{\cdot}28$). As is indicated in Figure 1, two of the transparencies have simulated distances beyond the physical distances that were obtainable in the alley. If these two largest simulated distances are eliminated from the analysis, the average line of best fit for the transparency presentations is $D' = 0{\cdot}41D + 22{\cdot}5$ cm. with the slope of this line again significantly less at the 0·01 level ($t = 3{\cdot}90$) than the slope of the full-cue (alley) equation of 0·69 noted above. The equation of the straight line of best fit from the alley data indicates that the method of verbal report of distance was proportional to, but an underestimation of, physical distance under full-cue conditions of observation. The significantly greater slope of the equation determined from using the alley over that determined from using the transparencies indicates that the underestimation of the simulated distance occurring with respect to the familiar objects, viewed under otherwise reduced conditions, is too large to be accounted for by errors of underestimation involved in the method of verbal report. A valid underestimation of the simulated distance of the transparencies was present with this underestimation increasing with an increase in the magnitude of the simulated distance.

Discussion

Experiment 1 supports the view that both the retinal and the familiar size of known objects can influence perceived size under otherwise reduced conditions of observation. This conclusion is in agreement with the results from several previous experiments in which measurements of the perceived size of familiar objects were obtained under reduced conditions of observation. In one of these studies (Gogel and Newton, 1969) it was found, using the same transparencies as in the present study, that the observer gave off-sized responses similar to the computed off-sized responses shown in Table I. Both studies are in good agreement in showing the marked effect of retinal size on the perceived size of familiar objects under otherwise reduced conditions of observation. In another study in which pairs of transparencies of familiar objects were presented, the ratio of perceived size to familiar size again varied as a function of the ratio of the retinal size to the familiar size (see Gogel and Mertens, 1968, Table I). Finally, from two other studies (Gogel, 1969*a*; Gogel and Mertens, 1967) it is evident that retinal as well as familiar size was a determiner of the perceived size of a familiar object presented under otherwise reduced conditions. It is clear that the inability of familiar size to dominate perceived size under otherwise reduced conditions is well documented. It seems reasonable to expect that familiar size, as a learned cue based on a long and continuing experience, would be a decisive factor in perceived size. Its limitations in this respect introduce doubt concerning the effectiveness of learned determiners of perceived spatial extent.

The results from the present study support the conclusion that the contribution of retinal size to perceived size (via the specific distance tendency) can result in the perception of off-sized objects. It is further asserted that the perception that the objects are off-sized is used to make inferential judgements of distance. From Experiment 2, this process results in systematic errors in distance estimation as a function of simulated distance. Several hypotheses can be considered as to the source of these errors. (1) The errors could be the result of a tendency for the distance judgement to be displaced in the direction of the perceived distance resulting from the specific distance tendency. In this case the distance judgements would have a perceptual as well as an inferential aspect. (2) The errors could be the result of not perceiving the off-sized objects as much off-sized as would be expected from the ratio of retinal size to familiar size. (3) The errors could be the result of some general inferential error in converting the perception of off-sized objects to a judgement of distance. It should be noted that the first two possibilities would require that the judged distances of perceptually under-sized objects would be underestimated but also that the judged distance of perceptually over-sized objects would be overestimated (relative to their simulated distances). But, it will be noted also that the perceptually under-sized objects had large values of S_F and D_F while the perceptually over-sized objects had small values of S_F and D_F. The first two possibilities listed above would require, therefore, that the slope of the transparency data of Figure 1 would be greater for the objects of small simulated distances than for the objects of large simulated distances. Perhaps this expectation is reflected in the tendency for the full-cue and the reduced condition curves of Figure 1, to be most similar for small simulated distances. Additional research is required, however, to distinguish between these three possibilities.

From this study it is obvious that the process underlying familiar size as a cue to distance is complex. A description of this process consistent with the data is as follows. When the observer is presented with a familiar object in an otherwise uniform field of view, with all other distance cues removed, he tends to perceive the object at a distance determined by the specific distance tendency, with the consequence that the perceived size of the object is determined by this perceived distance and the visual angle of the object. When asked to give a verbal report of the apparent size of the familiar object, the observer responds on the basis of either (1) the perceived size resulting from the retinal size, (2) the known size of the object resulting from its familiar characteristics, or (3) as is most likely, some compromise between (1) and (2). Since the perceived size resulting from the above process is frequently different from the known, familiar size, the observer usually perceives the object as being an off-sized object. When asked to report the apparent distance of this perceptually off-sized object the observer responds inferentially, with the reported distance varying systematically as a function of the amount of the off-sized appearance. The smaller the object appears to be compared with a normal-sized object of that type, the greater is the reported distance. However, this distance report frequently is not correct with respect to the simulated distance, i.e. not correct with respect to the distance at which the familiar object would have to be placed to subtend that particular retinal size. Simulated distances beyond about 10 ft. tended to be underestimated in this study even when the usual tendency to underestimate with verbal reports is taken into consideration. Also the magnitude of the underestimation tended to increase with increasing values of the simulated distance.

The distance reports obtained in this study were successive reports of egocentric distance. It has been found that judgements of egocentric distance from familiar size can differ depending upon whether the results from first or from successive presentations are being considered (Gogel, 1968*a*, 1969*a*; Gogel and Mertens, 1967). It needs to be determined whether the curve of D' vs. D from the transparency presentations would be similar to Figure 1 if the data from only first presentations were considered. Because of the relatively large number of different familiar objects used in the present study only a small number of observers received the same familiar object first. The data from the present study, therefore, is not appropriate for the problem of experimentally investigating the effect of first and successive presentations on the transparency data of Figure 1.

In the present study the different simulated distances were produced with different familiar objects. Different simulated distances can also be produced by using different retinal sizes of the same familiar object. It is possible that the curve of Figure 1 resulting from the presentations of the transparencies is independent of which of these two methods is used in producing the simulated distances. This possibility is supported by a study in which it was found that the difference between the reported and the simulated distance of a playing card increased as the visual angle of the playing card decreased (Gogel and Mertens, 1967). It is likely that the error in estimating simulated distances is a function of the discrepancy between perceived size and familiar size regardless of whether the familiar objects presented are the same or different objects.

A number of conclusions can be drawn from the present study and from related research (Gogel, 1968*b*, 1969*a*; Gogel and Mertens, 1967; Gogel and Newton,

1969) regarding judgements of the perceived size and distance of a familiar object. (1) Retinal size (via the specific distance tendency) contributes importantly to the determination of the perceived size of familiar objects under otherwise reduced conditions of observation. (2) The report of the distance of a familiar object under otherwise reduced conditions of observation probably is inferentially determined despite the instructions to make perceptual responses. (3) The resulting judgements of distance involve systematic errors as a function of the simulated distance of the familiar object. (4) Since an inferential process probably is involved in these distance judgements considerable caution is required in using familiar objects to test the interaction of perceived size and perceived distance as expressed in the size-distance invariance hypothesis. Additional research is needed to determine the reason for the systematic errors in judging the distance of the familiar objects when only familiar size is a possible cue to distance.

This study was supported by Grant NGR 05-010-010 from the National Aeronautics and Space Administration. I wish to thank Michael Koslow and Robert Newton for their help in collecting and analysing the data of the study.

References

CARLSON, V. R. (1960). Overestimation in size-constancy judgments. *Am. J. Psychol.* **73**, 199–213.

CARLSON, V. R. (1962*a*). Size-constancy judgments and perceptual compromise. *J. exp. Psychol.* **63**, 68–73.

CARLSON, V. R. (1962*b*). Underestimation in size-constancy judgments. *Am. J. Psychol.* **75**, 462–5.

EPSTEIN, W. (1963). Attitudes of judgment and the size-distance invariance hypothesis. *J. exp. Psychol.* **66**, 78–83.

GOGEL, W. C. (1968*a*). The effect of set on perceived egocentric distance. *Acta Psychol.* **28**, 283–92.

GOGEL, W. C. (1968*b*). The measurement of perceived size and distance. In *Contributions to Sensory Physiology*, Vol. III. NEFF, W. D. (Ed.). New York: Academic Press.

GOGEL, W. C. (1969*a*). The absolute and relative size cue to distance. *Am. J. Psychol.*, in press.

GOGEL, W. C. (1969*b*). The sensing of retinal size. *Vision Res.*, **9**, 8–24.

GOGEL, W. C., HARTMAN, B. O. and HARKER, G. S. (1957). The retinal size of a familiar object as a determiner of apparent distance. *Psychol. Monogr.* **71** (13, Whole No. 442), 1–16.

GOGEL, W. C. and MERTENS, H. W. (1966). Problems in Depth Perception: Perceived size and distance of familiar objects. Federal Aviation Agency Report AM66–22.

GOGEL, W. C. and MERTENS, H. W. (1967). Perceived size and distance of familiar objects. *Percept. Mot. Skills* **25**, 213–25.

GOGEL, W. C. and MERTENS, H. W. (1968). Perceived depth between familiar objects. *J. exp. Psychol.* **77**, 206–11.

GOGEL, W. C. and NEWTON, R. E. (1969). Perception of off-sized objects. *Percept. Psychophys.* **5**, 7–9.

JENKIN, H. and HYMAN, R. (1959). Attitude and distance-estimation as variables in size matching. *Am. J. Psychol.* **72**, 68–76.

Received 12 *May* 1969

Q. Jl exp. Psychol. (1969) **21**, 248–255

THE INFLUENCE OF ACOUSTIC SIMILARITY ON SERIAL RECALL OF LETTER SEQUENCES

NANCY S. ANDERSON

University of Maryland, College Park

Acoustic similarity is known to impair short-term memory (STM) for letter sequences. The present series of experiments investigated the effects of acoustic similarity on long-term retention. In the first experiment, subjects were asked to learn one of two lists of 8 letters, the letters being either of high or low acoustic similarity. Lists were visually presented for three trials, with subjects responding after each trial. Then subjects participated in an immediate memory task for digits which lasted for 20 min. Finally, subjects tried to recall the list of letters they had learned previously. Lists having items of high acoustic similarity were more difficult to recall on the first trial, but were better recalled on the delayed retention test. In a second experiment, groups of subjects were again asked to learn one of two lists of 8 letters differing in acoustic similarity, using different orders of the letters used previously. The procedures were identical except that in two groups, a STM task for digits intervened between the presentation and test of the letters. This intervening task minimized the effects of STM and eliminated the differences in retention found previously. In a third experiment, better long-term retention for material having high acoustic similarity was also obtained when subjects used a backward recall procedure. In the last experiment 14 item lists were learned to a criterion of two correct trials, and retention was tested after each trial and at a delay of 20 min. and 23 hr. No effect of acoustic similarity was found and little retention loss occurred. These results suggest that reducing the STM component by introducing a STM control or by lengthening the list caused the effect of acoustic similarity to disappear.

Introduction

Several experimenters (Conrad, 1964; Wickelgren, 1965) have shown that STM for sequences of letters is impaired when the letters are chosen from a set which is high in acoustic similarity. The interest in STM has increased and several investigators have postulated that both short-term memory (STM) and long-term memory (LTM) lie on a single continuum and that retention losses can be explained in both by concepts of interference theory (Keppel and Underwood, 1962; Melton, 1963; Postman, 1964). Baddeley (1966) has suggested that many of the difficulties in interpretation of similarities in STM and LTM may be due to the different techniques used to study retention in the two areas.

Several experimenters have taken phenomena known to influence LTM and have attempted to demonstrate the same effects in STM (Keppel and Underwood, 1962; Murdock, 1963). However, Baddeley (1966) chose to study a phenomenon known to influence STM, that of acoustic similarity, and studied its effect on LTM. In another study, Baddeley and Dale (1966) suggested "that STM may differ from LTM in relying more on acoustic cues and much less on the meaning of the material to be retained."

The series of experiments reported below used a similar approach to that of Baddeley (1966). Using letter sequences matched for predictability (frequency of occurrence in the language), the effects of the amount of acoustic similarity were investigated in serial recall for both STM and LTM.

Experiment 1

Method

Two lists of 8 consonants were made up, List H (QBDVCTGP) had an average probability of acoustic confusion of 0·73 (S.D. = 0·1) and List L (XNWKRYLH) had an average probability of acoustic confusion of 0·20 (S.D. = 0·1). Acoustic confusion or similarity was calculated following Conrad and Hull (1964) assuming the constant ratio-rule for subsets (Clarke, 1957). Both lists were arranged such that the average rank order of frequency of occurrence of letters as well as diagram frequency of occurrence of pairs of letters were approximately the same for both lists. Thus, predictability according to Baddeley, Conrad and Hull (1964) was matched across the two lists for the sequence as a whole and for each serial position. The average rank order predictability was 15·6, and the Spearman rho for diagram predictability = 0·93.

Housewives from the A.P.R.U. panel were tested in separate groups for each list of letters. Subjects were tested in groups of 8 or 10 and were instructed about the nature of the serial learning task consisting of three trials. Each of the 8 letters was projected on a screen at the rate of one every 2 sec. Subjects were given a ready signal before the first letter in each trial and a signal indicating the end of the sequence. Subjects wrote in an answer booklet as many of the letters as they could in the same order in which they were shown. As in Baddeley's (1966) and Conrad's (1964) studies, an alphabetic arrangement of the set of letters which comprised the list was visible throughout the experiment. This procedure was followed to maximize response availability throughout the experiment. To prevent the use of position on this display as a cue, four displays with different position arrangements were used. Subjects were allowed to guess a letter if they could not remember it, and were allowed 30 sec. to write in the booklet as much as they remembered. After the first trial, subjects turned over the page in the booklet and were not allowed to look back. This procedure was followed for three trials, after which they were tested on an intervening task (for 20 min.) involving immediate memory for digits. The procedure for this intervening task used a taped presentation of 12 digits at 1/sec. or 3/sec. with a post stimulus visual display card after each sequence presented with one item omitted, and subjects attempted to recall only a single digit. This procedure is a variation of the probe digit experiment reported by Waugh and Norman (1965).

After the intervening task, subjects tried to recall the eight letters they had learned previously in the appropriate order, on a single response sheet. Subjects were not told in advance about the subsequent retest, during which the relevant letters on a display were again made available.

Results

Performance for the 18 subjects each for List H and List L was scored in terms of items correct per position on each trial including the delayed retention, and the average percentages of letters correct are shown in Figure 1. Comparisons between lists were conducted using non-parametric tests. The differences between H and L were significant on trial 1 and the delayed retention test ($P = 0·05$

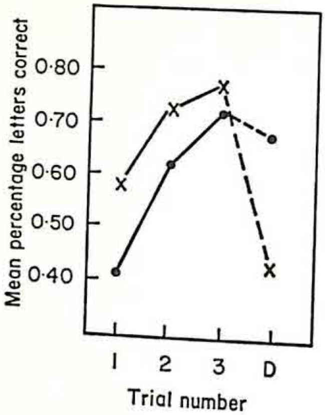


FIGURE 1. Learning and retention of 8-letter sequences as a function of acoustic similarity. (●), High; (×), low.

and $P < 0·01$) but on no other trials. The retention loss was significant only for List L ($P < 0·01$). These results show that on trial 1, the high acoustic similarity is detrimental to short-term memory supporting the work by Conrad (1964), but that these items once learned are retained significantly better than low similarity items over the 20-min. delay period. The fact that high acoustic similarity results in better retention supports the work of Baddeley (1966), who found similar results with words. Performance on the delayed retention test for List L was lower than it was on the first trial, suggesting that performance during learning might be dependent upon a STM component, and further experiments were designed to investigate this.

Experiment 2

Method

Two different serial ordered lists of eight letters were again conctructed. List H (VCQGDBPT) had an average acoustic similarity of 0·74 and List L (YRLXKHWN) had an average similarity of 0·20. Again the lists were balanced for predictability. (Average rank of 15·6, and Spearman rho = 0·97.) Separate groups of subjects were tested for each list of letters and for each of the three conditions of the experiment. Condition I of this experiment (recall only) was identical in procedure to Experiment 1. Condition II (STM prior to recall) of this experiment required that different groups of subjects recall 6 sequences of 8 random digits in a standard digit span paradigm. This STM procedure occurred between the presentation and recall of the 8 letters, and was the same as that used by Baddeley (1966) to control STM, and similar to that used by Glanzer and Cunitz in free recall (1966). That is, subjects saw a sequence of letters, then participated in an STM task for digits prior to recalling the letters. The effect of the digit span task was to reduce the influence of the STM component on the recall of the letters. In order to prevent subjects from writing the letters prior to instructions, special response sheets with two columns were used, and subjects were not told in which column to write the letters until after the STM task. Condition III of this experiment (STM post-recall) was conducted as a time control for Condition II. That is, the STM task for digits in Condition II takes time and the entire experiment is longer. To verify that the length of the experiment *per se* had no differential effect on recall, the STM task was introduced after recall of the letters, that is between recall and the next list presentation or trial. Because some retention loss between trials would occur, four trials were used. Other than these differences, all procedures of response availability, instructions, and the delay task were identical to those of Experiment 1. (Numbers of subjects learning the High and Low Lists were: Condition I, $N = 15, 14$; Condition II, $N = 15, 16$; Condition III, $N = 14, 14$.)

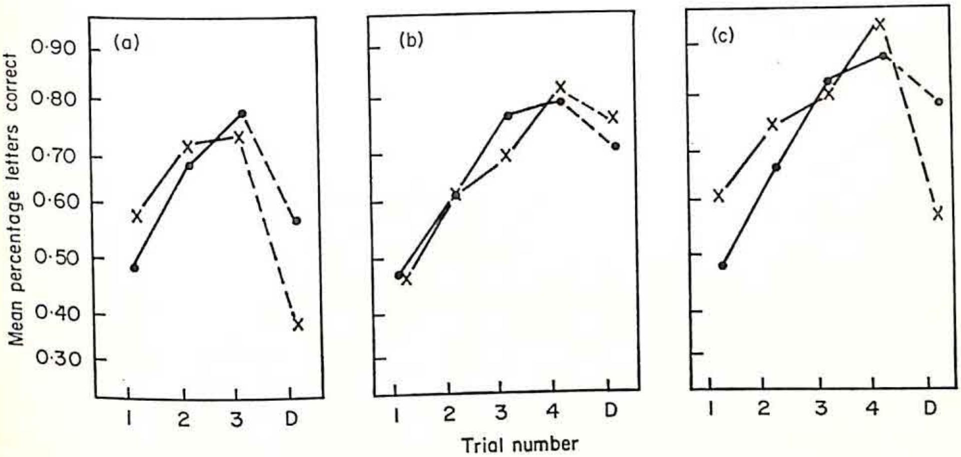


FIGURE 2. (a) Recall only; (b) STM prior to recall; (c) STM Post-recall. Learning and retention of 8-letter sequences as a function of acoustic similarity. (●), High; (×), low.

Results

Figure 2 shows the percentage of letters correct for each group of subjects for Lists L and H for each part of the experiment. The results of Condition I (recall only) are similar to those in Experiment 1, although the difference in Lists L and H on trial 1 was not statistically significant. The difference between Lists L and H on the delayed retention test was significant, and there were significant retention losses between trial 3 and delayed retention for both groups ($P < 0\cdot05$ in all cases). No differences were found on trials 2 and 3. In Condition II (STM prior to recall) no significant differences were found between lists, and in fact no trial comparison was as large as one standard deviation unit. The results of Condition III (STM post-recall) are essentially the same as those found in Condition I. The difference between Lists L and H on trial 1 was in the correct direction, but significant only at $P < 0\cdot10$, and the difference on the delayed retention test was significantly higher for List H than List L ($P < 0\cdot05$). Again, trials 2, 3 and 4 did not differ significantly. Both lists did show significant retention losses.

These results are in agreement with Baddeley's (1966), in that the STM task for digits removes the influence of the STM component during the letter learning task, and thus the differences found in Experiment 1 were removed. It appears then, that the differences obtained show that the acoustic similarity discourages a subject from using a strategy that relies on his STM. Whatever strategy he adopts to overcome these effects of acoustic similarity does result in better LTM for these items. Since List L has less acoustic similarity and is easier to process in STM, subjects can use the STM component to boost performance on each trial. However, after a delay of 20 min. when the STM component has presumably disappeared, the performance on List L is even lower than it was on trial 1. (In the delayed retention tests (recall only and STM post-recall) in this experiment and

Experiment 1, subjects who had learned List L looked shocked when the experimenter asked them to recall the list unlike the subjects who had learned List H who started immediately to write their responses.) The poor retention on the delayed test and the disappearance of these differences with the role of STM reduced, supports the notion that a strategy difference occurs for subjects who learn list H (compared to those who learn List L) where they must overcome the deleterious effects of acoustic similarity on STM.

Experiment 3

Method

Further studies were conducted using University of Maryland college students to replicate these results. Since the average age of college students is lower than the housewives on the A.P.R.U. panel, and since the testing was conducted in larger groups, alternate subjects were instructed to recall the list forward or backwards in order to reduce copying. Subjects did not know in which order they were to write responses until after the list was exposed on trial 1 and they turned over the response booklet (labelled Backward or Forward). After the three learning trials and the 20 min. immediate memory task for digits, subjects were asked to recall the sequences in the order learned. The data for 5 subjects who wrote the delayed recall in a forward order when the learning data were in a backward order were eliminated. The procedures were used to gain some control over backward order of recall when responses were written. Procedures were otherwise the same as in Experiment 1. The list orders were those used in Experiment 2.

Results

Since a ceiling effect was found for the forward recall data, they are not presented here although they are similar to the data in Experiment 1. The backward recall data are shown in Figure 3. (List H, $N = 22$; List L, $N = 31$.) Although there were no significant differences between Lists L and H on trials 1, 2 and 3, the difference on the delayed test was significant ($P < 0{\cdot}01$). List H showed an average increment of 4 per cent compared to a decrement on List L of 24 per cent.

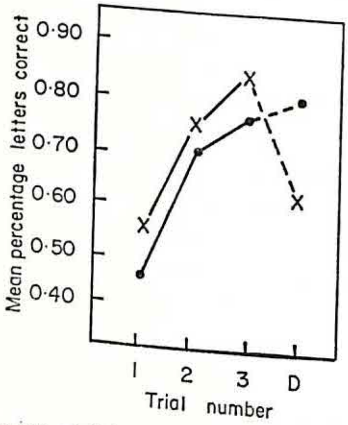


FIGURE 3. Learning and retention of 8-letter sequences (backwards) as a function of acoustic similarity. (●), High; (×), low.

There were small differences in Lists H and L on trial 1. In this case, the lack of a significant effect of acoustic similarity may be because backward recall is harder, and a different strategy is adopted for List L (which is not simply forward recall in STM) to overcome this difficulty.

Experiment 4

Method

A further study was conducted with University of Maryland students using 14-item lists with forward recall. Each of the lists, H and L, was a random arrangement of 7 consonants with the restriction that no two adjacent letters were the same. List H had an average acoustic similarity of 0·74 and List L of 0·13. The orders of each of the 14-letter lists were again matched for predictability (average rank predictability = 15·8 and Spearman rho = 0·97). The procedures were the same as in previous experiments, except that all subjects were tested individually, and were taken to a criterion of two perfectly correct trials before the 20-min. delay was introduced, and they were also tested 23 hr. after the short delay test. Subjects knew after the first session that they would be returning the next day.

Results

Examination of the average number of correct consonants on each trial showed no significant differences between lists, even on trial 1. The average number of consonants recalled on the delayed retention tests was: 20 min., High 12·9 (S.D.= 1·4) and Low 12·6 (S.D. = 2·1); 23 hr, High 12·2 (S.D. = 1·8) and Low 11·2 (S.D. = 3·6). The average number of trials to criterion was 11·2 (S.D. = 5·9) for the High similarity list and 11·7 (S.D. = 4·8) for the Low similarity list. There were indeed only small differences between lists and little loss in retention occurred.

These results indicated that for a longer list of 14 items which were learned to criterion, no differences between lists occurred during learning, and no significant differences occurred during the delayed retention tests. That no differences between lists occurred on the first trial seems to show that longer lists are harder to encode in STM. Furthermore, the lack of a retention loss after criterion was reached could be due to overlearning.

Discussion

The first three experiments are in agreement with Baddeley's (1966), and suggest that the strategy adopted by subjects for the acoustically similar list helps retention over the delayed interval. The poor LTM performance for the non-similar list is presumably due to the fact that the items are few in number and subjects rely on their STM during the learning trials. The use of the STM control prior to recall (Experiment 2, Condition II), made it difficult for subjects to rely on the STM component during learning, and the differences between the High and Low similar lists disappeared. Glanzer and Cunitz (1966) have used a similar procedure to isolate the role of STM in free recall, and Murdock (1963) has also demonstrated the interaction of STM and LTM.

It is to be noted that in both the present series of studies and Baddeley's (1966), the particular responses were made available to subjects during all the learning and retention tests. Thus, it is probably not a simple sound-like characteristic of the similar lists which would make them more available, since both response sets were made available to subjects. Whatever strategy was adopted to overcome the deleterious effects of acoustically similar material in STM, it is effective by trial 2 in the present studies. That is, although the higher acoustically similar lists were significantly poorer on trial 1, the poorer performances on these lists were not significantly different from the lower similarity lists on trials 2 and 3. However, in Baddeley's data, the acoustically similar list is significantly poorer than the control list throughout learning, and only on the retention test are the two lists not different, demonstrating less forgetting for the similar list. This difference between the present data and his data may be due to the fact that the lists of 8 letters were easier than his 10-word sequences.

The data from the last experiment support the STM-LTM component distinction, for in longer lists, the STM component is reduced. That is, the STM component on a long list is reduced on the first part of the list by the interpolated presentation of the last part of the list, and similarly, if subjects attempt to recall items in the order in which they were presented, then the recall of the first part of the list reduces the STM component for the last part of the list. This interpretation is consistent with Tulving and Arbuckle (1963) for paired associate learning, and the results of post-stimulus cuing for digits found by Anderson (1960). Since the STM component is reduced for the longer lists in Experiment 4, the performance differences between the similar and non-similar lists disappeared as they did with the STM control.

These experiments lend support to the distinction between STM and LTM, indicating that when performance on the first trial was degraded by the acoustic similarity on the STM component, subjects adopted a strategy which resulted in better retention over a delay than when first trial performance was boosted by an STM component. Thus, material difficult for STM may not be difficult for LTM. What particular strategy is adopted to overcome the deleterious effects on the STM component to code for LTM needs further investigation.

Part of this research was conducted while the author was an NSF Senior Postdoctoral Fellow visiting at the Applied Psychology Research Unit, Cambridge.

References

ANDERSON, N. S. (1960). Post-stimulus cuing in immediate memory. *J. exp. Psychol.* **60**, 216–21.

BADDELEY, A. D. (1966). The influence of acoustic and semantic similarity on long-term memory for word sequences. *Q. Jl exp. Psychol.* **18**, 302–9.

BADDELEY, A. D., CONRAD, R. and HULL, A. J. (1965). Predictability and immediate memory for consonant sequences. *Q. Jl exp. Psychol.* **17**, 175–7.

BADDELEY, A. D. and DALE, H. C. A. (1966). The effects of semantic similarity on retroactive interference in long- and short-term memory. *J. verb. Learn. verb. Behav.* **5**, 417–20.

CLARKE, F. R. (1957). Constant ratio rule for confusion matrices in speech communication. *J. acoust. Soc. Am.* **29**, 715–20.

CONRAD, R. (1964). Acoustic confusions in immediate memory. *Br. J. Psychol.* **55**, 75–84.

CONRAD, R. and HULL, A. J. (1964). Information, acoustic confusion and memory span. *Br. J. Psychol.* **55**, 429–32.

GLANZER, M. and CUNITZ, A. (1966). Two storage mechanisms in free recall. *J. verb. Learn. verb. Behav.* **5**, 351–60.

KEPPEL, G. and UNDERWOOD, B. J. (1962). Proactive inhibition in short-term retention of single items. *J. verb. Learn. verb. Behav.* **1**, 153–61.

MELTON, A. W. (1963). Implications of short-term memory for a general theory of memory. *J. verb. Learn. verb. Behav.* **2**, 1–21.

MURDOCK, B. B. (1963). Short-term memory and paired-associate learning. *J. verb. Learn. verb. Behav.* **2**, 320–8.

POSTMAN, L. (1964). Immediate memory and incidental learning. In MELTON A. W. (Ed.), *Categories of Human Learning*, pp. 145–201. New York: Academic Press.

TULVING, E. and ARBUCKLE, T. Y. (1963). Sources of intratrial interference in immediate recall of paired associates. *J. verb. Learn. verb. Behav.* **1**, 321–37.

WAUGH, N. C. and NORMAN, D. A. (1965). Primary memory. *Psychol. Rev.* **72**, 89–104.

WICKELGREN, W. A. (1965). Acoustic similarity and intrusion errors in short-term memory. *J. exp. Psychol.* **70**, 102–8.

Received 12 *May* 1969

Q. Jl exp. Psychol. (1969) **21**, 256–266

USING SEVERAL MODALITIES TO PERCEIVE ONE TEMPORAL PATTERN

STEPHEN HANDEL AND LOUIS BUFFARDI
Department of Psychology, Kansas State University

In two experiments, subjects identified temporal patterns. The patterns consisted of eight dichotomous (left-right) elements, e.g. LLRRLRLR, continuously repeated until the subject was able to identify the pattern. In Experiment 1, one pattern was presented either separately in the auditory, tactual, or visual modalities or one pattern was presented simultaneously in two modalities (compatible presentation). In Experiment 2, one pattern was simultaneously presented in two modalities, but the pattern was presented in one modality and the complement of the pattern (the complement of LLRRLRLR is RRLLRLRL) was presented in the second modality. Therefore, opposite spatial elements appeared in each modality (incompatible presentation).

The results indicated that the rate of pattern identification was the same for compatible and incompatible presentation. These methods produce better performance than individual modality presentation at fast presentation rates (2 elements/sec. and faster) although individual modality presentation was better at slower rates. This suggests that when a pattern is presented in two modalities, the pattern in each modality is integrated, not the particular spatial elements in each modality. Furthermore, the rate of pattern identification using individual modalities did not predict the difficulty using pairs of modalities. These results demonstrate the Gestalt nature of pattern perception; the pattern is perceptually salient and the performance of pairs of modalities depends on the inherent properties of the individual modalities.

Introduction

In two experiments, subjects organized and identified temporal patterns presented in the auditory, tactual or visual modalities and pairs of these modalities. The aim of the experiments was to compare perception using individual modality presentation with perception using two-modality presentation.

Previous work comparing individual modality presentation with two- or three-modality presentation of correlated stimuli has used simple stimuli and simple responses, including reaction-time experiments (Todd, 1914; Swink, 1966), detection experiments (Brown and Hopkins, 1967), and vigilance experiments (Buckner and McGrath, 1963). These workers concluded that increasing the number of modalities produces better performance. However, theoretical analysis and experimental research suggests that these tasks investigate simply "sensations" not "perception". Gibson (1966) argued that higher-order invariants (i.e. gradients and patterns) are the perceptual variables and Garner (1966) has summarized research indicating that a pattern and the number of its perceived alternates determines: (a) the judged goodness of a pattern; (b) the associations to a pattern; and (c) the ease of identifying a pattern. In accord with this view, the stimulus in the present experiment was a temporal pattern, the response was a description of the pattern, and subjects were free to organize the pattern in any way they wished.

Temporal patterns

The stimuli were sequences of dichotomous elements generated by a repeating pattern of eight dichotomous elements (i.e. LLRRLRLRLLRRLRLR). The subject observed the sequence passively and when ready, attempted to describe the pattern verbally. Two measures of performance were obtained: (a) the number of elements presented until the pattern was correctly identified; and (b) the manner in which the sequence was organized into a pattern, shown by the verbalized pattern description.

Previous work has shown: (a) regardless of the pattern formed by the first eight elements of the sequence, the pattern used to describe the sequence began with a series of identical elements (e.g. LLLL) or with a series of alternating elements (e.g. LRLR) (Royer and Garner, 1966, 1969; Handel and Buffardi, 1968). Therefore, the first element of a series of identical elements or the first element of a series of alternating elements are preferred start points; (b) a pattern and its complement, produced by exchanging the dichotomous elements (the complement of LLLRLRLR is RRRLRLRL), are perceptually equivalent, since both patterns were described identically and the rate of identification was the same (Royer and Garner, 1966); (c) starting the sequence at an actual preferred start point produced faster pattern identification than starting at a non-preferred start point for rates of presentation up to 4 elements/sec. (Garner and Gottwald, 1968); (d) patterns were organized similarly whether they were presented in the auditory or visual modalities, but the rate of identification depended on the modality used to present the pattern (Garner and Gottwald, 1968).

Pattern presentation

Three different methods of pattern presentation were used in this study.

Individual modality presentation. Here, patterns were presented in either the auditory, tactual, or visual modality. The dichotomous elements in each modality were identifiably different left-right stimuli. Although the dichotomous elements were different, localization was the dominant perceptual cue in each modality (also found by Handel and Buffardi, 1968; Royer and Garner, 1966). Therefore, the elements in each modality were different, but the pattern description of one pattern, in terms of the left-right sequence, was identical for each modality.

Compatible presentation. Here, the same pattern was presented simultaneously in two or three modalities; if a left element appeared in one modality, a left element simultaneously appeared in the second and third modalities. The spatial elements in each modality and the description of the pattern presented in each modality were identical.

Incompatible presentation. Here, a pattern was presented in one modality, and the complement of the pattern was presented simultaneously in the second modality. For example, if the pattern in one modality was LLRLLRLR, the pattern in the second modality would be RRLRRLRL; a left element in one modality was simultaneously presented with a right element in the second modality and vice versa. On the basis of spatial localization, which is the dominant characteristic, the elements presented in each modality were incompatible.

This presentation procedure represents stimulus-stimulus incompatibility as distinguished from stimulus-response incompatibility (see Fitts and Seeger, 1953). The patterns presented in each modality were identical, although the spatial elements and pattern descriptions were incompatible. The subject can use either modality to identify and describe the pattern, or the subject can use both modalities to identify the pattern in terms of series of identical elements or alternating elements (i.e. LLLLRLLR may be learned as 4121). As discussed previously, patterns are organized by these series. Then, the subject chooses one modality to describe the pattern.

Present experiment

The present report concerns two experiments comparing pattern identification and organization using individual modality presentation, compatible presentation and incompatible presentation.

General Procedure and Apparatus

Sensory elements

The dichotomous elements in each modality differed both by spatial position and by easily discriminated sensory differences.

The auditory stimuli consisted of two tones presented by loudspeakers placed about 6 ft. apart, 4 ft. from the subject. One tone was 1200 Hz and the second tone was 3000 Hz. The intensity of both tones was 65 dB sound pressure level.

The tactual stimuli were generated by a pair of vibrators (Sherrick, 1965). The subject held one vibrator in each hand and his hands were separated by 1 ft. One vibrator was powered by a 12 v, 60 Hz source, and the second vibrator was powered by a 6 v, 30 Hz source.

The visual stimuli consisted of two 1 in. panel lights mounted 1 ft. apart on a Masonite board placed vertically on a table. The board was placed at eye-level, 4 ft. in front of the subject. One panel light was red and the other green; the brightness of each light was 0·25 ft.-L.

Stimulus presentation

The stimuli were controlled by a tape reader (Western Union 1B) which activated a set of relays to switch the stimuli on and off. Toggle switches controlled the choice of modalities. Two timers (Hunter, Model 111C) were used to time the sequence. A counter recorded the number of elements presented.

The subject sat in a soundproof room. The experimenter observed the subject through a one way mirror and communicated with him by intercom.

Task

The subject observed the pattern without making any overt response until he thought he could identify the pattern. Then, the subject stopped the presentation by pressing a foot switch, and attempted to describe the pattern by verbally stating the left-right sequence. If the description was incorrect, the experimenter stated that the response was incorrect and restarted the pattern at the element where it was stopped. If the description was correct, the next pattern was presented. If the pattern had not been correctly identified after 560 elements of the pattern, the pattern was terminated and the next pattern was presented.

Experimental variables

Three rates of presentation were used: 1, 2 and 4 elements/sec. At each rate of presentation, each element appeared for the first third of the inter-element interval.

Two start points, determined from Royer and Garner (1966), were used for each pattern. The preferred start point (PSP) was an element often used by subjects to begin the pattern description. The non-preferred start point (NPSP) was an element rarely used by subjects to begin the pattern description.

Patterns

The 10 patterns used, shown in Table I, are the same ones used by Garner and Gottwald (1968). Each pattern and its complement were presented equally often.

TABLE I

Ten patterns used in Experiments 1 and 2

	Start point	
Pattern	Preferred	Non-preferred
1	LLLLLRRR	LLLRRRLL
2	LLLRLLRR	LRLLRRLL
3	RLLRLLLL	RLLLLRLL
4	LLRLLRLR	RLLRLRLL
5	RLRLRLLL	LLRLRLRL
6	RLRLRRLL	LRRLLRLR
7	RRLRLLLL	LLRRLRLL
8	RRLLRLLL	LLRRLLRL
9	LLLRLRRR	LRRRLLLR
10	LLLLRLRR	LRRLLLLR

The patterns are shown at the preferred and non-preferred start points.

Performance measures

Pattern identification. The number of elements presented before a pattern was correctly identified (delay) was the measure of rate of pattern identification. The median delay was found at every experimental condition. The delay, averaged across the 10 patterns, was calculated for every rate of presentation × start point × modality condition since the rank order of difficulty of patterns for each modality condition was approximately equal.

Pattern organization. The measure of correct pattern identification was the percentage of pattern descriptions beginning at the actual pattern starting point. The experimental variables which affect organization were start point and method of pattern presentation. Therefore, the percentages of descriptions were averaged across rates of presentation, patterns, and modalities for each start point and method of pattern presentation.

Experiment 1

In Experiment 1 patterns were presented using individual modality presentation or compatible presentation.

Method

Subjects. The subjects were 84 Kansas State University undergraduates. Each subject was paid for the task which took about 75 min.

Modalites. Seven modality conditions were used: the 3 individual modalities [audition (A), tactile (T), vision (V)], the 3 pairs of modalities [audition-tactile (A-T), audition-vision (A-V), tactile-vision (T-V)] and the single combination of all three modalities [audition-tactile-vision (A-T-V)].

Experimental design. The basic unit of the design was a block consisting of the ten basic patterns, each presented for a single trial. Five of the patterns started at a PSP, and 5 of the patterns started at a NPSP. In addition, 2 of the rates of presentation were used for 3 patterns apiece and the remaining rate of presentation was used for 4 patterns. All 10 trials in a block were presented in the same modality condition.

Every subject was presented three different blocks, so that summed across the 30 trials: (a) each pattern was presented once at each rate of presentation; and (b) each pattern started twice at one start point and once at the other start point. A different modality condition was used for each block. Therefore, each subject was presented 30 of the 420 possible combinations of 10 patterns, 3 rates of presentation, 2 start points, and 7 modalities; each subject was presented 3 of the 7 modality conditions and 10 of the 60 possible conditions for each modality.

The order of presentation of the patterns was counterbalanced in the following manner. Six different orders of 3 blocks were constructed. The sum of the first blocks of the 6 orders, 60 patterns in all, consisted of the 3 rate of presentation × 2 start point × 10 pattern conditions. Similarly, the sum of the second blocks and the sum of the third blocks also consisted of these conditions. Therefore, each condition appeared once during the first block of 10 trials, once during the second block, and once during the third block.

To form the different modality sets for the 3 blocks in an order, 14 different sets of 3 modality conditions were chosen from the 35 possible sets. Each of the 7 modality conditions occurred 6 times in the 14 sets. Every set of 3 modalities has 6 different orders so that each of the 6 orders of a set of 3 modalities was paired with 1 of the 6 pattern orders. This was done for all 14 sets of 3 modalities leading to 84 unique combinations of pattern orders and modality orders. A different subject was randomly assigned to each unique combination and all 84 subjects counterbalanced the order of presentation of the patterns and modality conditions. Every pattern × rate of presentation × start point × modality condition was presented to 6 subjects.

The subjects were pretrained on two patterns not used in the experiment. All modality and rate of presentation conditions were demonstrated. During the experimental session, the subject was told the rate of presentation and modality condition before each pattern was presented.

Results

Pattern identification. The delay, averaged across the 10 patterns, at every rate of presentation × start point × modality condition is shown in Table II.

TABLE II

Individual modality and compatible presentation

	Rate of presentation (elements/sec.)						
	1		2		4		
	Start point		Start point		Start point		
Modality	Preferred	Non-preferred	Preferred	Non-preferred	Preferred	Non-preferred	Average delay
A	25	37	36	66	131	156	75
T	26	41	40	68	133	253	94
V	24	36	38	63	255	278	116
A-T	27	39	41	61	135	191	82
A-V	24	34	32	57	91	112	58
T-V	23	42	44	74	145	236	94
A-T-V	23	37	31	67	128	160	74
Average	25	38	37	65	145	198	

The average number of elements presented until the patterns were correctly identified.

The effects of rate of presentation and start point were straightforward†; increasing the rate of presentation produced significantly greater delay in pattern identification and, at every rate of presentation, starting the pattern at the NPSP produced significantly greater delay in pattern identification. These results replicate those of Garner and Gottwald (1968) and generalize them to the T modality and combinations of modalities.

There was no consistent difference in rate of pattern identification using individual modalities since the rank order of difficulty changed as the rate of presentation increased. At the slow rate of presentation there was little difference among modalities although patterns presented in the V modality were slightly easier to identify. At the fast rates of presentation, patterns presented in the A modality were easiest to identify, those in the T modality next, and patterns presented in the V modality were hardest to identify (observations also made by Garner and Gottwald, 1968). In addition, when patterns started at PSP, patterns presented in the T modality were easy to identify (similar to the rate of identification of patterns presented in the A modality), but when patterns started at NPSP, patterns presented in the T modality were hard to identify (similar to the rate of identification of patterns presented in the V modality).

There were significant differences in rate of pattern identification among pairs of modalities. Pattern presentation using the A-V combination produced the best performance; the results in the A-T combination were intermediate; and the T-V combination produced the poorest performance. Furthermore, only the A-V combination produced significantly better pattern identification than its component modalities. At all 6 rate of presentation × start point conditions, the A-V combination was easier than either the A or V modalities. The rate of pattern identification using the A-T or the T-V combinations was intermediate to the rate of their component modalities.

---

† Analyses of variance were carried out on the logs of the median delay until pattern identification for both Experiments 1 and 2. All effects discussed were significant at least at the 0·05 level.

Pattern organization. When patterns started at PSP, the patterns were usually described beginning at the PSP whether they were presented in individual modalities (89 per cent) or in combinations of modalities (89 per cent).

When patterns started at NPSP, and were presented in individual modalities, 19 per cent of the descriptions began at the NPSP. When patterns started at NPSP, and were presented in combinations of modalities, 12 per cent of the organizations began at the NPSP. A further analysis indicated that for 27 per cent of the patterns presented in individual modalities, the NPSP was the element most often used to begin the pattern description. For the 15 per cent of patterns presented in combinations of modalities, the NPSP was the element most often used to begin the pattern description. Both measures differ significantly ($P < 0{\cdot}05$, z-test) between individual modality presentation and two- or three-modality presentation. Therefore, when patterns started at NPSP, and were presented in two- or three-modalities, subjects were more likely to describe the pattern starting at PSP than when patterns were presented in one modality (66 per cent vs. 58 per cent).

Discussion

There is no simple means of predicting from the rate of pattern identification using individual modalities which pair of modalities will produce faster pattern identification than its component modalities. Although patterns presented in the A or T modality were easiest to identify, patterns presented in the A-V combination were easier than any other pair or triplet of modalities. In addition, only the A-V combination produced faster pattern identification than its component modalities. The rate of pattern identification in a pair of modalities, therefore, seems to be independent of the rate of pattern identification of its component modalities.

These results suggest that the A and V modalities complement each other when they are combined; in some way each modality presents "information" not present in the other modality. Vision and audition are normally thought of as being different kinds of senses; V is simultaneous and spatial, A is sequential and temporal (Liberman, Cooper, Harris, MacNeilage and Studdert-Kennedy, 1967, Murdock, 1968). The rate of pattern identification using pairs of modalities may depend on how the inherent properties of the modalities mesh; the perceptual characteristics of pairs of modalities are not simply the summation of the characteristics of the individual modalities.

Experiment 2

In this experiment, patterns were presented using incompatible presentation.

Method

Subjects. The subjects were 36 Kansas State University undergraduates and high school seniors; they were paid for an experimental session of about 75 min. No subject from the previous experiment was used.

Modalities. The 3 pairs of modalities from Experiment 1 were used.

Experimental design. The experimental design counterbalanced both the order of presentation of the modalities and the order of presentation of the patterns.

The method of counterbalancing the patterns was exactly the same as used in Experiment 1. The 3 pairs of modalities have 6 different orders and all 6 orders of the modalities were paired with each of the 6 pattern orders resulting in 36 unique combinations. One subject was assigned to each combination. Every pattern × rate of presentation × start point × modality condition was presented to 6 subjects. For every pair of modalities, each modality presented the pattern and its complement equally often.

The subjects were pretrained on two patterns not used in the experiment. All modality and rate of presentation conditions were demonstrated. During the experimental session, the subject was told the rate of presentation and modality condition before each pattern was presented.

Results

Pattern identification. The delay, averaged across the 10 patterns, at every rate of presentation × start point × modality condition is shown in Table III.

TABLE III

Incompatbile presentation

	Rate of presentation (Elements/sec.) 1 Start point		2 Start point		4 Start point		
Modality	Preferred	Non-preferred	Preferred	Non-preferred	Preferred	Non-preferred	Average delay
A-T	26	38	39	57	127	214	84
A-V	23	47	32	65	111	158	73
T-V	30	45	48	60	126	220	88
Average	26	43	40	61	121	197	

The average number of elements presented until the patterns were correctly identified.

Increasing the rate of presentation produced significantly greater delay in pattern identification. At every rate of presentation, starting a pattern at a NPSP produced significantly greater delay than starting the pattern at a PSP.

There were significant differences among pairs of modalities; the A-V combination produced the best pattern identification, the A-T combination was intermediate and the T-V combination produced the poorest identification.

Pattern organization. When patterns started at PSP, the pattern was usually described beginning at the PSP (86 per cent). When patterns started at NPSP, 16 per cent of the descriptions began at the NPSP. For 23 per cent of the patterns, the NPSP was the element most often used to begin the pattern description.

Since the subject can correctly identify the sequence by describing the pattern presented in one modality or the complement presented in the second modality, it is possible to determine if modality preferences exist. The pattern presented in the

A modality was used for 90 per cent of the descriptions using the A-V combination and for 88 per cent of the descriptions using the A-T combination. The pattern in the T modality was used for 75 per cent of the descriptions using the T-V combination.

Discussion

Pattern identification when using compatible or incompatible presentation was essentially the same; the effects of rate of presentation and start point were identical and the pairs of modalities in both methods had the identical rank order of difficulty. The subjects demonstrated modality preferences. The A modality was preferred to the T or V modalities, and the T modality was preferred to the V modality.

General Discussion

There was no difference in rate of pattern identification between compatible and incompatible presentation; both methods produced faster pattern identification than presentation in a single modality. The rank order of the modality conditions, measured by their delays averaged across experimental conditions, was: A-V (58), A-V Incompatible (73), A (75), A-T (82), A-T Incompatible (83), T-V Incompatible (88), T (93), T-V (94), V (116). Averaged across modalities, the delay using incompatible presentation was 78 elements, the delay using incompatible presentation was 81 elements, and the delay using individual modality presentation was 94 elements.†

Further analysis shows the conditions in which two-modality presentation produced superior identification than individual modality presentation. The 6 rate of presentation × start point conditions were split into 2 sets of 3 conditions apiece. The first set of conditions was the easiest three conditions: 1 element/sec.-PSP, 1 element/sec.-NPSP, 2 elements/sec.-PSP. For all 3 of these conditions, individual modality presentation was at least as easy as either method of simultaneous presentation. The average delay using individual modality presentation was 34 elements, the average delay using compatible presentation was 34 elements, and the average delay using incompatible presentation was 36 elements.

The second set was the three hardest conditions: 2 elements/sec.-NPSP, 4 elements/sec.-PSP, 4 elements/sec.-NPSP. For all 3 of these conditions, both methods of simultaneous presentation were always easier than individual modality presentation. The average delay using compatible presentation was 123 elements, the average delay using incompatible presentation was 126 elements, and the average delay using individual modality presentation was 156 elements.

The superior pattern identification in pairs of modalities may simply result from subjects concentrating only on the 1 modality which produced the fastest pattern identification. However, out of 9 ordinal comparisons (3 pairs of modalities × 3 experimental conditions), compatible presentation of pairs of modalities produced faster identification than both component modalities 5 times, intermediate identification 2 times, and slower identification than both component modalities 2 times. Incompatible presentation produced faster identification than both component modalities 6 times, intermediate identification 3 times, and never produced slower identification than both component modalities. Therefore, for these 3 conditions, both methods of simultaneous presentation produced faster pattern identification than presentation in the best (in rate of pattern identification) component modality of a pair of modalities.

---

† The superior performance of two-modality presentation has been replicated in subsequent work (Lewis, 1968). The aim of that study was to determine the effect of practice on pattern perception. When subjects were naive, incompatible presentation produced significantly faster pattern identification than either compatible presentation or individual modality presentation. In addition, compatible presentation produced slightly better pattern identification than individual modality presentation. After about 10 hr. of practice, these differences disappeared and all methods were of equal difficulty.

Our results clearly support the distinction made by Garner and Gottwald (1968) between pattern learning and pattern perception. At slow rates of presentation, up to 2 elements/sec., patterns are learned as a sequence of elements; a process which is unintegrated, derived and active. At faster rates of presentation, patterns are perceived as a whole; a process which is phenomenally integrated, immediate and passive.

The major difference between the 2 sets of experimental conditions is rate of presentation. When patterns are learned (first set), presentation of the pattern in 2 modalities does not improve pattern identification. But when patterns are perceived (second set), compatible and incompatible presentation produce faster pattern identification. The similarity in rate of pattern identification using either method of simultaneous presentation leads to the conclusion that what is being integrated is the *pattern*, not the spatial elements, and not the pattern descriptions.

The distinction between learning and perception resolves the apparent anomalous result that stimulus-stimulus incompatibility improves the rate of pattern identification. At slow rates of presentation, the pattern is learned as a series of elements and the stimuli are the spatial elements; the spatial elements are indeed incompatible and the incompatibility produces the poorest performance. At faster rates of presentation, the pattern is perceived and learned as a unified whole. The *patterns* are the stimuli; the stimuli are not the spatial elements. Since the pattern in each modality is identical, the stimuli are not incompatible, and the subject can use the identical pattern in each modality to identify the pattern.

This research was supported in part by the Bureau of General Research, Kansas State University, and in part by NASA University Sustaining Grant NSG-692.

We thank Dr. C. E. Sherrick for loaning two vibrators.

References

Brown, A. E. and Hopkins, H. K. (1967). Interaction of the auditory and visual sensory modalities. *J. acoust. Soc. Am.* **41**, 1–7.

Buckner, D. N. and McGrath, J. J. (1963). A comparison of performances on single and dual sensory mode vigilance tasks. In Buckner, D. N. and J. J. McGrath (Eds.), *Vigilance, A Symposium.* New York: McGraw-Hill.

Fitts, P. M. and Seeger, C. M. (1953). S-R compatibility: Spatial characteristics of stimulus and response codes. *J. exp. Psychol.* **46**, 199–210.

Garner, W. R. (1966). To perceive is to know. *Am. Psychol.* **21**, 11–9.

Garner, W. R. and Gottwald, R. L. (1968). The perception and learning of temporal patterns. *Q. Jl exp. Psychol.* **20**, 97–109.

GIBSON, J. J. (1966). *The Senses Considered as Perceptual Systems.* Boston: Houghton-Mifflin.

HANDEL, S. and BUFFARDI, L. (1968). Pattern perception: integrating information presented in two modalities. *Science, N.Y.* **162**, 1026–8.

LEWIS, W. E. (1968). The effect of practice on the identification and organization of temporal patterns. Unpublished Master's dissertation, Kansas State University.

LIBERMAN, A. M., COOPER, F. S., HARRIS, K. S., MACNEILAGE, P. F. and STUDDERT-KENNEDY, M. (1967). Some observations on a model for speech perception. In WATHEN-DUNN, S. (Ed.), *Models for the Perception of Speech and Visual Form.* Cambridge: M.I.T. Press.

MURDOCK, B. B., Jr. (1968). Modality effects in short-term memory: storage or retrieval? *J. exp. Psychol.* **77**, 79–86.

ROYER, F. L. and GARNER, W. R. (1966). Response uncertainty and perceptual difficulty of auditory patterns. *Perception Psychophys.* **1**, 41–7.

ROYER, F. L. and GARNER, W. R. (1969). Perceptual organization of nine element auditory temporal patterns. *Perception Psychophys*, in press.

SHERRICK, C. E. (1965). Simple electromechanical vibration transducer. *Rev. scient. Instrum.* **36**, 1893–4.

SWINK, J. R. (1966). Intersensory comparisons of reaction time using an electro-pulse tactile stimulus. *Human Factors* **8**, 143–5.

TODD, J. W. (1912). Reaction to multiple stimuli. *Arch. Psychol. N.Y.*, Whole No. 25.

Received 12 *May* 1969

Q. Jl exp. Psychol. (1969) 21,267–271

IMPROVEMENT IN MEMORY WITH TIME

FELICIA A. HUPPERT† AND J. A. DEUTSCH
University of California, San Diego

Groups of rats which had incompletely learnt to escape from shock into the lit arm of a Y-maze, were tested for retention of the habit at various times after training. These times ranged from 30 min. to 17 days. Retention was found to be poor from 30 min. to 5 days, showed a marked improvement at 7 to 10 days and a subsequent decrement after 14 days. These results parallel those previously found for the differential susceptibility to cholinergic drugs of habits of different age. They provide further evidence that memory strength increases with time up to a point, and then gradually decreases.

Introduction

It has been noted (Deutsch, Hamburg and Dahl, 1966; Deutsch and Rocklin, 1967; Hamburg, 1967; Wiener and Deutsch, 1968) that memories which have been stored for different lengths of time are differentially sensitive to treatment with cholinergic drugs. Such differences in susceptibility span days or weeks, rather than minutes or hours (as has been reported for most other treatments; Deutsch, 1962; McGaugh, 1966). It has been found that at times when control rats forget a Y-maze habit (28 days after training), rats injected with anitcholinesterase show facilitation of memory for the habit (Deutsch and Leibowitz, 1967; Wiener and Deutsch, 1968). However, at times when control rats remember the habit (up to 28 days), there is no simple relationship between the behaviour of controls and the behaviour of rats injected with the drug. When administered 30 min. after training, the drug produces a slight impairment; 1 and 3 days after training the drug has no effect, and 7 days after training there is a severe impairment (Deutsch *et al.*, 1966; Hamburg, 1967). Since the drug results are believed to be a reflection of the underlying strength of the memory, it might be argued that one should be able to see differences in the strength of a memory in an untreated group of animals at various times after training. The fact that such a result was not found in previous studies may be attributable to the training procedure, which involved training animals to a criterion of 10 consecutive correct responses. Using this criterion, the performance of control animals tends to be perfect during retest at least for the first 2 to 3 weeks after training. Thus a "ceiling" effect may have been operating in the previous studies.

In the present study we attempted to circumvent the "ceiling" effect by undertraining the rats. Previous studies have shown that rats take about 36 trials to reach the criterion of 10 consecutive correct trials. In this study, we gave the rats only 15 trials initially, and then measured the number of trials taken to reach the criterion of 10 consecutive correct trials at various times after initial training.

---

† Present address: Psychological Laboratory, University of Cambridge.

Method

Subjects

One hundred and sixteen Sprague-Dawley rats, weighing 300 to 350 g. at the time of training were obtained from a local source (Miller, Vista) and maintained in group cages on an *ad lib.* diet.

Apparatus and procedure

Subjects were trained and tested in an automated Y-maze. Each arm was 19 in. long, 5 in. wide and 6 in. high. At the end of each arm was a plexi-glass shield behind which was a 60 w incandescent lamp supplied with 65 v. a.c. when lit. A 0·75 mA shock could be applied to the grid floor of the maze. In the course of the experiment, shock was continually applied to the centre of the maze, extending for 6 in. into each alley. The terminal 13 in. of each alley received shock when that alley was unlit, but was shock-free when the alley was lit. Only one alley was lit at any one time. Illumination of lamps and switching of shock were automatically controlled by a digibit logic system. Entry by the subject of more than 6 in. into a lit alley interrupted a photo-cell beam which activated the logic system. The terminal 13 in. of a lit alley, which was shock-free, remained so for a 20 sec. interval after entry by the subject. Immediately following this period, the light was turned off in the shock-free arm, shock was applied to that arm, and another alley was lit and made shock-free. The sequence of alley illumination followed a semi-random order pre-wired into the logic system.

In the training phase each rat received 15 trials in the maze. A correct response was scored when the subject's initial choice on a trial was the lit alley and the subject escaped into the shock-free terminal section of the alley. An incorrect response was scored when the subject entered more than 6 in. into an unlit alley before escaping into the lit alley. Rats were removed from the maze upon completion of the 15 trials and returned to their home cage, where they remained for varying intervals before being tested. In the test phase, rats were given further trials in the maze until they attained a criterion of 10 consecutive correct responses. Rats taking more than 70 trials to reach criterion were discarded, as were those which had reached criterion during the 15 trials of the training phase.

An experiment and a replication were run. There were 50 rats in each study allocated at random to one of 5 groups, each group representing a different training-testing interval. In the first experiment, the intervals between training and testing were 30 min., 1, 3, 7 and 14 days. In the replication, the intervals used were 1, 5, 7, 10 and 17 days.

In addition, a control group of 16 rats was run (control 0). These rats were simply trained to the criterion of 10 consecutive correct responses.

Results

Table I shows the mean, median and standard deviation of the number of trials taken to attain the learning criterion as a function of the training-testing interval. The comparability of the results of the two studies was analysed by means of two Mann–Whitney U-tests. There was no significant difference between the means of the two 1-day groups or between the means of the two 7-day groups. The data were therefore pooled for subsequent analysis.

As can be seen in Figure 1, there is a large decrease in trials to criterion when animals are tested 7 to 10 days after training. Included in Figure 1 are the results

TABLE I

Mean, median and standard deviation of trials to criterion as a function of the interval between training and testing

		Time between training and testing				
		30 min.	1 day	3 days	7 days	14 days
Experiment 1:	Mean	26·4	26·0	23·6	10·2	27·8
	Median	22·5	19·5	20·0	1·5	27·0
	S.D.	19·2	20·7	12·2	14·1	10·0
		1 day	5 days	7 days	10 days	17 days
Replication:	Mean	19·4	25·1	14·1	9·9	33·9
	Median	11·5	24·0	9·5	8·5	29·5
	S.D.	26·3	20·2	16·8	9·7	12·2

for control 0, where rats were simply trained to criterion without prior training. This group took a mean of 35·9 trials to reach criterion.

A Kruskal–Wallis analysis of variance by ranks (Seigel, 1956, pp. 184–93) indicated that there was a significant overall effect due to training-testing interval ($H = 20{\cdot}55$, $d.f. = 7$, $P < 0{\cdot}01$). A series of specific comparisons was then carried out using Mann–Whitney U-tests. These showed that: (a) the 7 and 10 day groups were not significantly different from one another; (b) the 30 min., 1, 3, 5, 14 and 17 day groups were not significantly different from one another; and (c) the 7 and 10 day groups were significantly different from each of the other groups, with the *f*-value 0·05, or better, in each case with the single exception that the difference between the 10 day group and the 1 day group fell just short of significance.

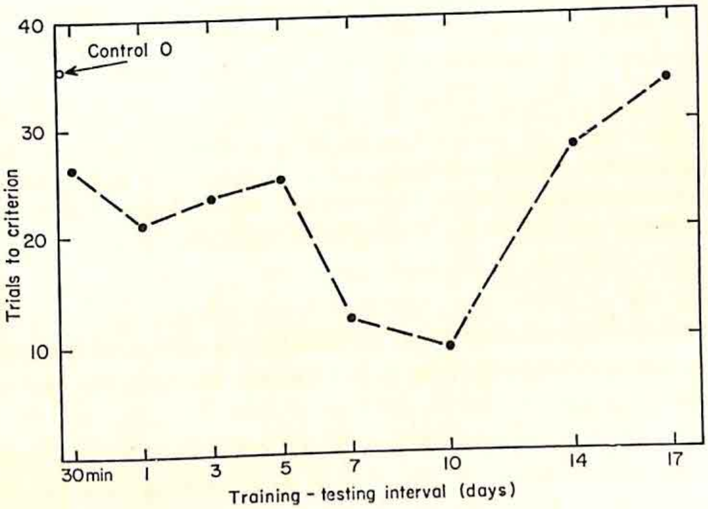


FIGURE 1. The mean number of trials taken by undertrained rats to relearn a Y-maze habit, as a function of the training-testing interval. (The data from Experiment 1 and the replication have been pooled.) Control 0 represents the score of naive animals trained to criterion on the same task.

Discussion

The results indicate that memory for an undertrained habit is stronger 7 to 10 days after training than it is at shorter post-training intervals, and that by 17 days the habit is almost completely forgotten.

Previous data have demonstrated that susceptibility to anticholinesterase drugs such as DFP and physostigmine is far greater for older memories (7 to 14 days) than for more recent memories (up to 5 days); i.e. older memories are completely blocked by these drugs, whereas newer memories are either unimpaired or only partially impaired. As one would expect block with an anticholinesterase only when the normal release of the transmitter acetylcholine is very high, this finding has been interpreted as showing that transmission across cholinergic synapses increases with time after learning (Deutsch and Deutsch, 1966). As the peak of blockage in these experiments occurred somewhere between 7 to 14 days after training, it was inferred that this interval represented the peak of memory strength. The present study has shown that such a peak is evident at roughly the expected time, when purely behavioural tests are made in the same situation. We conclude that memory strength, at least for this task, at first increases with time and then decreases.

It has been suggested that the present results could be attributed to a fear incubation effect rather than to changes in memory strength. For instance, Kamin (1963) found that performance of a partially learned shock avoidance habit shows an initial decrement and then gradually improves with time since initial training; McMichael (1966) found that conditioned suppression of a bar-press response increases with time since training. These effects are generally attributed to an increase of fear with time. To fit a fear interpretation to the present findings, we would have to say that fear is *decreasing* with time, since performance in our discrimination task improves with time and since we use an escape rather than an avoidance procedure and an increase of fear could only have a deleterious effect. If our results were attributable to a decrease in fear leading to improved performance with time, then only the eventual deterioration of the habit by 17 days would reflect alterations in memory strength, probably due to forgetting.

However, an identical time course of memory susceptibility to cholinergic drugs has been demonstrated in an escape situation and in an appetitive situation (Wiener and Deutsch, 1968). Changes in the level of fear, therefore, cannot explain all the phenomena obtained with the use of drugs. As the behavioural findings reported here parallel closely, and were predicted from, the pharmacological studies, it seems reasonable to suppose that their basis is the same and not due to motivational factors. Indeed, one may well ask whether there is adequate evidence to show that performance changes with time after training are ever caused by fear incubation. It is entirely possible that the changes attributed to fear incubation are actually memory changes.

McMichael (1966) has attempted to demonstrate that it is indeed fear which changes with time, by showing that the learning of a new habit is impaired at a time when the fear level is believed to be highest, i.e., when bar-pressing is maximally suppressed (4 hr. after training). Rats are first given unavoidable shock in the black compartment of a two compartment box, then at various times after this experience, replaced in the black compartment with shock switched off, and allowed to run into the neutral, white compartment. The 4-hr. group spend considerably less time in the white compartment than do the other groups, a finding which McMichael accepts as further evidence that fear is at a maximum at this time.

However, in the absence of any additional data, one could more naturally interpret this result as indicating that the 4-hr. group is *less* fearful than the others. Thus the possibility remains open that demonstrations of improved avoidance behaviour with time since training may reflect alterations in memory strength rather than fear incubation. What is needed are studies employing our paradigm but using an appetitively motivated task.

Another outcome of our experiment is of some interest. In an earlier study (Deutsch and Leibowitz, 1967), control rats trained to a criterion of 10 consecutive correct responses had forgotten the habit by 28 days. In our experiment the same task was incompletely learned, and it was found that the habit was forgotten by 17 days, although the peak of memory strength still occurred at the same time. This finding is evidence of a direct relationship between the initial strength of a habit and the length of time for which the habit is remembered.

References

DEUTSCH, J. A. (1962). Higher nervous function: the physiological bases of memory. *Ann. Rev. Physiol.* **24**, 259–86.

DEUTSCH, J. A. and DEUTSCH, D. (1966). *Physiological Psychology*. Homewood, Illinois: Dorsey Press.

DEUTSCH, J. A., HAMBURG, M. D. and DAHL, M. (1966). Anticholinesterase-induced amnesia and its temporal aspects. *Science, N.Y.* **151**, 221–3.

DEUTSCH, J. A. and LEIBOWITZ, S. F. (1966). Amnesia or reversal of forgetting by anticholinesterase, depending simply on time of injection. *Science, N.Y.* **153**, 1017–8.

DEUTSCH, J. A. and ROCKLIN, K. W. (1967). Amnesia induced by scopolamine and its temporal variations. *Nature, Lond.* **216**, 89–90.

HAMBURG, M. D. (1967). A retrograde amnesia produced by intraperitoneal injections of physostigmine. *Science, N.Y.* **156**, 973–4.

KAMIN, L. J. (1963). Retention of an incompletely learned avoidance response; some further analyses. *J. comp. Physiol. Psychol.* **56**, 713–8.

MCGAUGH, J. L. (1966). Time-dependent processes in memory storage. *Science, N.Y.* **153**, 1351–8.

MCMICHAEL, J. S. (1966). Incubation of anxiety and instrumental behaviour. *J. comp. physiol. Psychol.* **61**, 208–11.

SIEGEL, S. (1956). *Non-parametric Statistics for the Behavioural Sciences*. New York: McGraw-Hill.

WIENER, N. and DEUTSCH, J. A. (1968). The temporal aspects of anticholinergic and anticholinesterase induced amnesia for an appetitive habit. *J. comp. Physiol. Psychol.* **66**, 613–7.

Received 15 *May* 1969

Q. Jl exp. Psychol. (1969) **21**, 272–279

SOME EVIDENCE FOR PARALLEL COMPARISONS IN A LETTER RECOGNITION TASK

GILLIAN COHEN

Institute of Experimental Psychology, University of Oxford

Subjects were asked to judge successively presented letter trigrams "same" or "different." The different stimuli were divided into four groups; not confusable (NC), acoustically confusable (AC), visually confusable (VC), and both visually and acoustically confusable (VC & AC). Reaction times (RT) were lengthened only by the double confusability. It is argued that comparisons are normally made in both channels, so that confusability in a single channel has no effect since the alternative channel is unimpaired. RTs are only increased when both channels are slowed down. A further experiment confirms this interpretation. When the situation is manipulated so that only the visual channel is employed, the VC group shows the same increase as the VC & AC group.

Introduction

In visual recognition a stimulus is compared with a standard which is stored in memory, and classified as "same" or "different." In a letter recognition task it is conceivable that the matching process could be either acoustic (a name-name comparison), or visual (a shape-shape comparison). It is well established that acoustic recoding of visually presented letters commonly takes place in short term memory (Conrad, 1964; Wickelgren, 1965), but recognition experiments so far (Sternberg, 1967; Chase and Posner, 1965) point to a visual rather than an acoustic comparison process. Of course, acoustic matching cannot be entirely independent of visual information, since visual identification must precede naming, but the two systems present separate and distinct alternatives at the comparison stage, and can be separately manipulated (Posner, Boies, Eichelman and Taylor, 1969). The present experiments aim to test the possibility that letter recognition might be accomplished by both visual and acoustic matching operating in parallel. This possibility was suggested by a previous experiment (Cohen, 1969) in which subjects were required to make same/different judgements on two successively presented letter arrays. When the second array (the comparison stimulus) was different, it was either visually or acoustically confusable with the first array (the memory stimulus). However no significant difference between the two kinds of confusability was found. Obviously this result could occur if different subjects employed different strategies, the trends cancelling out in the pooled data; but in this case no individual subject showed a significant difference between the two

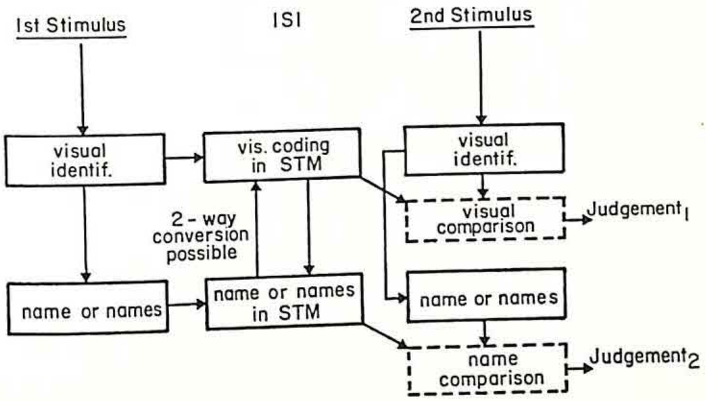


FIGURE 1. A model for recognition with visual and acoustic comparisons operating in parallel.

conditions. This finding gives rise to the following hypotheses which are tested in the present experiment:

(1) that the main effect of confusability occurs at the comparison or matching stage, rather than during identification;

(2) that if matching were *either* visual *or* acoustic, then confusability of the relevant kind should either slow reaction times or increase errors;

(3) that if both matching procedures were employed concurrently a single type of confusability would produce little or no effect, since one channel could still operate unimpaired;

(4) that if both visual and acoustic confusability are present simultaneously, then both matching procedures would be hampered and reaction times would be slower, and error incidence increased.

A parallel matching procedure also constitutes an explanation of the subject's ability to detect and correct his own errors (Rabbitt, 1968*a*). As the errors appear to represent fast premature responses, the mistake must be detected as a result of additional or alternative processing. A single comparison, or serial comparison model cannot explain the speed with which error correction follows the error response. If both visual and acoustic methods of recognition were employed in parallel, an error based on one of these matches could be detected as soon as the other procedure was completed.

Experiment 1

Method

Subjects. Six subjects completed the experiment. They were graduates and undergraduates of both sexes. All had previous experience as subjects in a similar experiment. They were paid at an hourly rate. In addition a pay-off system was instituted whereby subjects received 6d. extra for every RT below 350 msec., and lost 6d. for each error.

Materials. The stimuli consisted of consonant trigrams drawn from the set h s f p b d q c t. The letters were printed from lower-case Letraset, Type Face Plantin, 24 point, on 6 × 4 white cards, and subtended a visual angle of 5·29° horizontally, and 0·66° vertically. On 36 of the trials the second stimulus (CS—the comparison stimulus) was the same as the first

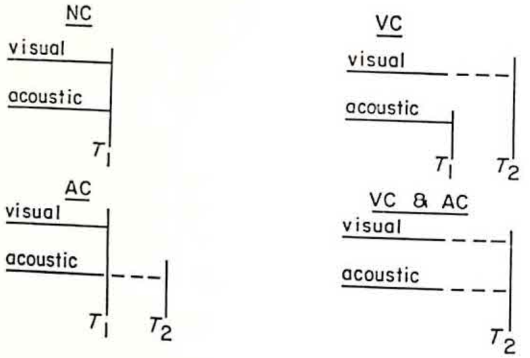


FIGURE 2. The time at which a judgement is available is only affected when both kinds of confusability are present.

stimulus (MS—the memory stimulus). On 36 trials it was different, in which case one letter of the CS trigram differed from the MS, and the different letter might occur in any position. The 36 "different" trials were further sub-divided into four types:

(1) *NC—not confusable.* In this group the substitutions were c for q and t for q, or the other way round. That is, the different letter was neither visually nor acoustically confusable with the one it replaced.

(2) *AC—acoustically confusable.* Here the substitutions were p for c and p for t. These letters were found to be acoustically confusable in the work of Conrad and Wickelgren (op. cit.)

(3) *VC—visually confusable.* The substitutions were b for q and p for q. These letters are the same shape at different rotations and were found to be visually confusing in some studies by Kolers (in press).

(4) *VC & AC—both visually and acoustically confusing.* In this group the substitutions were b for d and b for p.

In each group the substitutions occurred in both directions. The remaining background letters were equated for occurrence in all four groups.

Procedure. The cards were viewed through a Cambridge tachistoscope. The first stimulus card (MS) was exposed for 2 sec. After an ISI of 5 sec. during which the subject continued to fixate the stimulus location on a blank field, the second stimulus card (CS) was exposed following a verbal warning signal. The subject was instructed to respond "same" or "different" by pressing the appropriately labelled switch. The position of the two response switches was counterbalanced for hand across subjects. Half pressed the "same" switch with the dominant hand, and the "different" switch with the non-dominant hand; half had the opposite arrangement, The inter-trial interval was 10 sec., and the subject rested for 2 min. every 25 trials. Each subject had a practice run of 20 trials with a special practice set of materials, and then went on to the experimental trials which lasted about 45 min.

Design. The order of the trials was arranged so that the sequence of "same" and "different" trials varied randomly with the constraint that no sequence of one kind was longer than three trials. The four kinds of "different" trial (NC, VC, AC and VC & AC) were randomly distributed throughout the series with the constraint that they were balanced for position in the "same" and "different" sequences. For example, a "different" trial could occur following a "same" trial, following a "different" trial, or following two similar trials. Since response alternation and response repetition affect RT (Rabbitt, 1968*b*), it was thought necessary to balance for run-position. In order to offset practice and fatigue effects half the subjects received the trials in the reverse order.

Scoring. The time from the onset of the CS until the subject terminated the display by pressing a response switch was measured by a Decatron timer, accurate to within 10 msec. Both errors and RTs were recorded. The subject was told his RT at each trial.

Results

The mean RTs for the 6 subjects are shown in Table I. Error scores were excluded from the data. The scores for the four kinds of "different" judgement were tested by two-way analysis of variance, subjects × type of confusion, and found to be significantly different, $F = 10{\cdot}902$, $d.f.\ 3, 15$, $P < 0{\cdot}01$. A Newman–Keuls test modified for use with samples of unequal sizes (Winer, 1962) was performed to find which of the four conditions differed significantly from which of the others. The results showed that while VC & AC differs with $P < 0{\cdot}01$ from NC, AC and VC, these do not differ significantly from each other.

TABLE I

Mean RTs (msec.) for the four kinds of "different" judgements and for "same" judgement

Different				Same
NC	AC	VC	VC & AC	
436	437	452	529	396

The over-all error rate was 8·3 per cent and the error distribution is shown in Table II. The pay-off system was found to produce a higher than usual error rate, but a reduced RT variance. The error distribution follows the same trend as the RTs.

TABLE II

Error percentages for the four kinds of "different" judgements and for "same" judgements

Different				Same
NC	AC	VC	VC & AC	
7·4%	5·5%	9·3%	20·4%	6%

Discussion

The results show that both visual and acoustic confusability must be present if the RTs are to be significantly slowed, or the errors markedly increased. The presence of a single type of confusability does not raise the RTs or the errors above those of the non-confusable group. This result fits well with a parallel comparisons model. The effect of the four conditions on such a model is represented in Figure 2. Assuming that the comparison process is slowed by introducing confusability, the diagram shows that a judgement will be available at time T_1 in the NC, VC and AC conditions. Only in the VC & AC condition, when both comparison channels

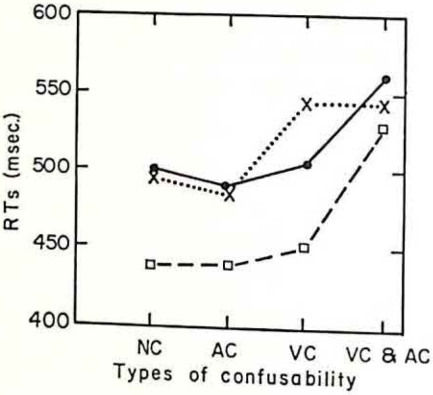


FIGURE 3. The results of Experiment I and of the two conditions in Experiment II. Reaction times as a function of different types of confusability.

are hampered, is the judgement delayed until time T_2. In the data $\bar{x}$VC is 15 msec. slower than $\bar{x}$ AC, which makes it appear that visual comparisons (operative when acoustic confusability has hindered acoustic comparison) are slightly faster. This difference, although not statistically significant, is in the expected direction since acoustic comparison requires an extra operation, name derivation.

The results of this experiment are clearly consistent with the parallel comparisons model, but some objections could be made to this interpretation. Of these, some can be ruled out by a closer inspection of the data.

(1) It may be objected that the substitution letters in the VC & AC group (b–d, and b–p) delay RTs more, not because they are both visually and acoustically confusable, but because they are *more* visually confusable than the substitution letters in the VC group (b–q and p–q). There is, however, no evidence to support this view. All these letters are the same shape at different rotations. In a recent paper Sekuler and Houlihan (1968) studied the discrimination of mirror images by adult humans to find out which rotations present most difficulty. They found that left-right mirror images (which would include p–q and b–d) are harder to discriminate than up-down mirror images (like p–b and b–q). In the present design both up-down and left-right rotations occur equally often in both the VC and the VC & AC groups, so that it is unlikely that the VC & AC group would be more visually confusable. It must be acknowledged that it was not possible to equate acoustic confusability between the AC and the VC & AC groups; p–c differ on two distinctive features, while p–t, b–p and b–d differ on only one, so that confusability might be marginally greater in the VC & AC group.

(2) It might be argued that the longer RTs in the VC & AC group reflect an increase in the identification time rather than an increase in the comparison time. If, for example, the letter "b" is harder to identify than other letters, a longer mean RT would result in a group which had a larger number of "b's" in the CS stimuli. To examine this possibility a *t*-test was performed to compare the mean RT for CSs with "b" within the VC & AC group with the mean RT for CSs with "b" in all other groups. The difference was significant at $P < 0{\cdot}005$. Further *t*-tests

showed no significant differences between different letters within the same group. In short, RTs in the data do not vary with the occurrence of any particular letter in the CS; they vary with the relationship between a letter in the CS and the one in the MS for which it is substituted. The longer RTs result from an increase in comparison time, not identification time.

(3) A further objection can be made based on the fact that all the substitution letters are not changed with equal frequency. The letter "p" is a member of three groups, the VC, the AC, and the VC & AC groups, and is therefore changed 9 times in all; but "d" is a member of only one group, the VC & AC, and is only changed 3 times. The other substitution letters c, t, q and b are all changed 6 times. On a response bias hypothesis the more often a given letter is changed, the faster a subject will detect such a change. A comparatively slow RT to a change on the letter "d" might therefore present an alternative explanation of the poorer performance in the VC & AC group. But comparison of the mean RTs for trials in which "p" is changed and trials in which "d" is changed show no significant difference. Within the VC & AC group the mean RT for "p" is 532 msec. and for "d" it is 536 msec. A response bias interpretation of the results is therefore not tenable.

Experiment II

A method was then sought to confirm the evidence for a recognition system operating more than one channel of comparison. If a single type of confusability fails to slow reaction times because an alternative channel or mode of recognition is still unhampered, then it follows that if the subject can be forced to employ one channel only, a single type of confusability in that channel should be effective in delaying judgement and slowing RTs. A method of manipulating the mode of comparison similar to that used by Posner (1967) was thought likely to be successful. The same materials were used, but the stimuli were to be judged under two different criteria. The original list was expanded so that an additional one third of the CSs were in upper case letters. The subject was then asked in one condition to judge the stimuli the same if they had the same names, i.e. to make an acoustic match. In this condition two kinds of pair were "same" (hsb = hsb and hsb = HSB), and two kinds were "different" (hsb $\neq$ hsp and hsb $\neq$ HSP). In the other condition the subject was asked to judge the stimuli "same" only if they were visually identical, i.e. to make a visual match such that only hsb = hsb was "same" and hsb $\neq$ HSB, hsb $\neq$ hsp and hsb $\neq$ HSP were all "different." In this way it might result that when an acoustic match was required, acoustic confusions slowed reactions; and with a visual criterion the visual confusions should have more effect.

Method

The method was as in Experiment I. Six new subjects attended for two sessions, working through one condition at each session. Half the subjects did the visual matching first and half did the acoustic matching first.

Results

Only the scores from trials with the lower-case CSs from the original list were used in the analysis. The mean RTs for the 6 subjects in both conditions are

TABLE III

Mean RTs (msec.) for the four kinds of "different" judgements and for "same" judgements under acoustic and visual criteria in Experiment II

	Different				Same
	NC	AC	VC	VC & AC	
Acoustic matching	502	488	509	558	469
Visual matching	490	485	539	539	439

shown in Table III. Two-way analyses of variance were carried out for each condition separately, subjects × type of confusability. Because the response entropy is not the same, the two conditions cannot be compared with each other.

In the name-matching task $F = 6{\cdot}1$, with *d.f.* 3, 15, which is significant at $P < 0{\cdot}01$. A Newman–Keuls test showed that the VC & AC mean differed significantly from the other three, which did not differ from each other.

In the visual matching task the data did not fulfil the requirements of the F test, since the variance differed markedly from subject to subject, but the less powerful Friedman two-way analysis of variance (Siegel, 1956) gave a result significant at $P < 0{\cdot}001$. A Newman–Keuls test on all the means showed that VC & AC and VC differed significantly from NC and from AC, but these pairs did not differ from each other. The level of significance was $P < 0{\cdot}01$.

Discussion

The results of Experiment II show that the visual matching instructions succeeded in manipulating the channel of comparison. Because the visual channel is employed alone in this task, visual confusability has raised reaction times to the same mean score as the double, visual and acoustic confusability. The subject is obliged to employ the visual channel of comparison alone, since acoustic comparison gives a conflicting result with upper-case stimuli. Hence a single type of confusability in the relevant channel, the visual one, raises RTs as much as double confusability in both visual and acoustic channels.

In the name matching task, however, the results parallel those of Experiment I, as Figure 3 shows. The scores are higher over-all because the introduction of upper case stimuli has complicated the task, but only the double confusability group differs significantly from the other three. It is apparent that the subject has not employed the acoustic channel alone in this task; but this finding is not surprising since visual identification must precede name derivation which would make it difficult to exclude visual comparison altogether.

The dual comparison model advanced here may well be of limited generality. Where, for instance, a stimulus has a single name, but considerable visual complexity, acoustic comparison may be optimal; and where verbal labels are very cumbersome it may be more economical to compare stimuli along visual dimensions. Also there is good reason to suppose that a dual comparison model is a very simplified one. In a less restricted situation matching multidimensional stimuli requires multiple comparisons. Whether these comparisons are made serially, or in parallel may vary with the nature of the task and the degree of practice. Although some studies such as Egeth's (1966) have indicated a serial processing model, the experiments presented here are more consistent with a parallel one.

This work was carried out while the author was in receipt of a Medical Research Council Scholarship.

References

CHASE, W. G. and POSNER, M. I. (1965). The effect of auditory and visual confusability on visual and memory search tasks. Paper read at Midwestern Psychological Association Meeting.
COHEN, G. (1969). The effect of codability of the stimulus on recognition reaction times. *Br. J. Psychol.* **60**, 25–29.
CONRAD, R. (1964). Acoustic confusion and immediate memory. *Br. J. Psychol.* **55**, 75–84.
EGETH, H. E. (1966). Parallel versus serial processes in multi-dimensional stimulus discrimination. *Percept. Psychophysics* **1**, 246–52.
KOLERS, P. A. (in press). Three stages of reading. In LEVIN, H. and WILLIAMS, J. (Eds.), *Basic Studies on Reading*. New York: Harper.
POSNER, M. I. (1967). Chronometric analysis of classification. *Psychol. Rev.* **74**, 392–409.
POSNER, M. I., BOIES, S. J., EICHELMAN, W. H. and TAYLOR, R. L. (1969). Retention of visual and name codes of single letters. *J. exp. Psychol. Monogr.* **79**, No. 1, Part 2.
RABBITT, P. M. A. (1968*a*). Three kinds of error-signalling responses in a serial choice task. *Q. Jl exp. Psychol.* **20**, 179–88.
RABBITT, P. M. A. (1968*b*). Repetition effects and signal classification strategies in serial choice-response tasks. *Q. Jl exp. Psychol.* **20**, 232–40.
SEKULER, R. W. and HOULIHAN, K. (1968). Discrimination of mirror images: choice time analysis of human adult performance. *Q. Jl exp. Psychol.* **20**, 204–7.
SIEGEL, S. (1956). *Non-parametric Statistics*. New York: McGraw-Hill.
STERNBERG, S. (1967). Two operations in character-recognition: some evidence from RT measurements. *Percept. Psychophysics* **2**, 45–53.
WICKELGREN, W. A. (1965). Acoustic similarity and intrusion errors in short-term memory. *J. exp. Psychol.* **70**, 102–8.
WINER, B. J. (1962). *Statistical Principles in Experimental Design*. New York. McGraw-Hill.

Received 22 *May* 1969

Q. Jl. exp. Psychol. (1969) **21**, 280–290

SYNTACTIC STRUCTURE MODIFIES ATTENTION DURING SPEECH PERCEPTION AND RECOGNITION

KENNETH ABRAMS
University of Michigan
AND
THOMAS G. BEVER
Rockefeller University

The present paper demonstrates the interaction of syntactic structure and speech perception with a response task which minimizes the effects of memory: reaction time (RT) to clicks during sentences. (1) In 12-word *unfamiliar sentences* each with two clauses, RT is relatively slow overall to clicks located at the end of the first clause but decreases as a function of clause length. Clicks at the beginning of the second clause are not affected by length of the preceding clause. (2) In *familiar sentences*, RT is relatively fast to clicks located at the end of a clause while RT to clicks at the beginning of clauses is relatively unaffected by familiarity. (3) RT is *not* fastest overall to clicks located between clauses either in novel or familiar sentences. (4) As in previous studies, the subject's subsequent judgement of the location of the click tone are towards the clause break. (5) We could find no systematic interaction between RT and subjective click location. Findings (1) to (3) are consistent with the view that perceptual processing alternates between attending to all external stimuli and developing an internal representation of the stimuli. Finding (3) is in conflict with an "information channel" view of immediate attention to speech, which would predict high sensory attention to non-speech stimuli between clauses. However, findings (4) and (5) indicate that the channel view of perception may be correct for that perceptual processing which occurs after the immediate organization of the speech stimulus into major segments.

Introduction

Recent studies have shown that subjects make systematic errors in locating a non-speech stimulus superimposed on speech. For example, when subjects hear a click during a sentence their subsequent estimates of the actual click location are biased toward the point between clauses. This finding has been interpreted as evidence for the view that listeners actively use grammar to impose clause structure on the speech stimulus as they hear it. In this paper we demonstrate this view with a response task that minimizes the possible influence of memory strategies and response biases. We also show that a "channel model" of attention applies to speech processing only after the speech signal has been segmented into perceptual major units.

Grammatical analysis divides sentences into major segments. For example, in sentence (1) there are two clauses,

(1) Peter ran quite fast but we caught him in the narrow valley.

"(Peter ran quite fast)" and "(but we caught him in the narrow alley)." The behavioural relevance of this kind of formal segmentation has been demonstrated in a number of experiments. Using a technique developed by Ladefoged and Broadbent (1960), Fodor and Bever (1965) found that in the dichotic presentation of a click and a spoken sentence, subjects recalled the location of the click as displaced toward the major phrase boundaries between clauses. For example, a click actually located in "fast" or in "but" in (1) was reported as occurring between "fast" and "but." These results have been extended using sentence materials which investigate intonational effects (Garrett, Bever and Fodor, 1966) the effect of sequence probability (Bever, Lackner and Stolz, 1969) and the relative effect of surface and underlying syntactic structure (Bever, Lackner and Kirk, 1969). The conclusion from these studies is that the immediate segmentation of stimulus sentences occurs at points between groups of phrases which are bound together by basic syntactic relations (e.g. the relations "actor, action, object, modifier").

In the click displacement experiments, subjects report the location of the click shortly *after* hearing (and writing down) the sentence. This allows not only instantaneous attentional strategies, but also delayed encoding strategies, and subsequent response strategies to affect the reported click location. Therefore, the previous studies show that the formal syntactic structure of a sentence plays a psychological role in immediate attention, in perceptual analysis, in organizing the response, or in some combination of these three processes.

Our concern in the present research was to investigate how the formal segmentation of speech affects *immediate* attention as measured by reaction time to clicks presented during the sentences. Reaction time to clicks is quite fast (100 to 300 msec.) so that any systematic effects of syntactic phrase structure on reaction time would occur within a syllable of the actual click presentation. Therefore, any systematic effects on reaction time could not be due to slightly delayed operations in the analysis of the sentence or to post-stimulus response strategies. The claim in previous accounts of click location errors has been that the points between clauses represent points of minimum attention to speech and maximum attention to other stimuli. Thus our primary hypothesis was that clicks objectively occurring *in* clause breaks should receive faster reaction times than clicks in any other location.

Experiment 1

Method

Nine 12-word sentences with two clauses each were recorded with subdued (but normal) intonation. These sentences had a clause break after the fifth, sixth or seventh word (see Appendix). Three tape copies of the original sentences were spliced into three different experimental orders such that sentences which were among the first third of one experimental order were among the middle and last third of the other two orders. In each sentence a single pulse was located on the second track of a stereophonic tape in one of three positions relative to the major clause break; in one copy it was placed in the word *before* the break, in one copy *after* the clause break and in one copy it was placed *in* the clause break. (The two words preceding and following the break were monosyllables.) Across the three experimental orders, a given sentence had pulses located in each of three positions relative to the clause break. In each experimental order there were three sentences with pulses before, three sentences with pulses in, and three sentences with pulses after the break; the sentences were presented so that each pulse position was represented once in each third of each experimental order. Thus, the experimental design was balanced across the three variables of clause break position, pulse position relative to the clause break and experimental order. Three normally-spoken 12-word practice sentences with a pulse preceded each experimental order of nine.

The pulse in each stimulus triggered a 25 msec. capacitor discharge click. The click was approximately equal in intensity to the loudest speech sound in the stimulus sentences. Triggering the click also started an electronic timer which was stopped when the subject closed a telegraph key.

Subjects were 12 M.I.T. undergraduates who volunteered their time. The subject was instructed to hold the fingers of his right hand on the key and to close it as soon as he heard the click. After responding to 20 isolated practice clicks, the subject then heard one of the three groups of stimulus sentences. He was told that he would hear a sentence and one click during each sentence. To insure high attention to the reaction time task, the subject was paid 25 cents for fast reaction time responses. (A "fast response" was defined without the subject's knowledge as any response faster than the fastest trial on the five preceding trials.) To insure attention to the sentence, the subject had to write out each sentence immediately after hearing it. The sentence was presented to the left ear and the click to the right ear.

After presentation of the first order of 12 sentences (three practice and nine experimental sentences, hereafter called "presentation 1") ,the subject read a book silently for 5 min. He was then presented with another order of the 12 sentences ("presentation 2"). After another 5 min. of silent reading, the subject was presented with the third order of the 12 sentences ("presentation 3"). Subjects were run individually, and the six possible orders of the three experimental groups of sentences were evenly balanced. Thus, every subject responded to every sentence three times, each time with the click in a different location relative to the clause break.

Results

There are two findings:

(1) The main hypothesis was not confirmed; clicks located in major clause breaks did *not* receive the fastest reaction times.

(2) Clicks located before clause breaks had longer reaction times than those located after clause break.

Table I

Mean reaction times to clicks before, in and after points between clauses, in first, second and third presentation of sentences (Experiment I sentences spoken naturally)

	1st presentation	2nd presentation	3rd presentation
Before-clause break	243	202	172
In-clause break	230	184	169
After-clause break	216	189	179

Table I presents the mean reaction time (RT) for each of the click positions in each of the presentation groups. The prediction that in-break clicks would be responded to fastest was not confirmed overall. This was tested first by comparing the average RT for each subject of *in-break* with *before-break* and with *after-break* clicks for all presentation groups together ($P < 0{\cdot}7$ one-tail by Wilcoxon matched-pairs signed ranks test on 12 subjects for each comparison. In this paper all significance measures are by a Wilcoxon matched-pairs signed ranks test across subjects or sentences, unless otherwise indicated). For the separate presentation groups the main prediction was not confirmed except for the difference between before-break clicks and in-break clicks in presentation 2 ($P < 0{\cdot}03$ one-tail).

Clicks preceding breaks were responded to more slowly than clicks following breaks. This was confirmed over all presentation groups considered together ($P < 0{\cdot}05$ two-tail). There was an experience effect: the slowness of RT's to before-break was confirmed for presentation groups 1 ($P < 0{\cdot}05$) and 2 ($P < 0{\cdot}025$) and was not confirmed for presentation 3.

Subjects' RT might have become faster as a function of how much of the sentence had passed; they might have noticed that the sentences were all about the same length, and could have adopted a strategy of increasing expectation for the click as they heard the sentence. If this were true, the relative slowness of RT to before-break clicks might not have been due to the syntactic structure itself, but rather to the fact that before-break clicks on the average appeared one word earlier than after-break clicks. To test this we calculated a separate average RT for each subject for each presentation order, of the responses to the sentences with the clause break after the fifth, sixth and seventh word. None of the differences were significant (by a Wilcoxon test on those subjects who contributed scorable responses in all syntactic positions in every presentation group).

If RT is affected by ongoing perceptual analysis, the results of this experiment suggest that the original view of the nature of the displacement of clicks in speech was not correct. Although before-break clicks receive slower RT than after-break clicks, in-break clicks do not receive the fastest RT. This implies that it is not the case that sensory attention to non-speech material is highest between clauses.

The failure to confirm our main hypothesis and the relative slowness of before-break clicks in Experiment I might be due to acoustic cues of structure such as intonation and pauses rather than due to the active use of structure in sentence perception, since the sentences in Experiment I were spoken with normal intonation. Experiment II was designed to eliminate the possibility of this kind of artifact. Subjects were also asked to locate the clicks after hearing each sentence so that we could compare the processes of click location and reaction to clicks.

Experiment II

Method

The three groups of 12 sentences used in Experiment I were reconstructed artificially. We first recorded the individual words separately in a random list. The appropriate words were then spliced together to produce sentences free from any systematic prosody cues. The placement of the pulses and constitution of the three experimental groups were identical to Experiment I.

Thirty right-handed, undergraduate native speakers of English were run in individual sessions. The subject responded by releasing a telegraph key (instead of closing it as in Experiment I) to 20 trials of a click heard in the right ear. As further practice, the subject then responded to clicks in the right ear with five normally-spoken 12-word sentences presented to the left ear. After this, the experimental sequence was exactly the same as in Experiment I: each subject received three groups of 12 sentences separated by 5 min. of silent reading. Through the experiment, the subject wrote down the sentences after responding to the click, and indicated where in the sentence the click had occurred. The subject was not given a reward for relatively fast responses.

Results

There are three findings (see Table II):

(1) The main hypothesis was not confirmed. In-break clicks did not receive the fastest RT.

(2) (a) RT to clicks at the ends of clauses was slower than to clicks at the beginning of clauses in presentations 1 and 2 and the opposite effect occurred in presentation 3.

(b) Within presentation 1, RT to clicks at the ends of clauses increased with the length of the clause, while RT to clicks at the beginnings of clauses was unaffected by the length of the preceding clause.

(3) (a) Post-stimulus click location responses replicated previous findings: Subjects tended to locate clicks in the break in all sentences and all presentations.

(b) There was no evidence that the mechanisms affecting reaction time and click location are the same.

TABLE II

Mean reaction times to clicks before, in and after points between clauses, in first, second and third presentation of sentences (*Experiment II sentences artificially constructed by tape-splicing*)

	1st presentation	2nd presentation	3rd presentation
Before clause break	248	207	180
In-clause break	239	201	193
After-clause break	234	194	196

The prediction that in-break clicks should have the fastest RT was not confirmed for any comparison in any presentation order: in-break responses are neither faster than before-break responses, nor are they faster than after-break responses for any presentation order.

As in Experiment I, the presentation order had a considerable effect on the extent to which before-break clicks are responded to more slowly than after-break clicks. In presentation 1, this effect was strongly confirmed ($P < 0{\cdot}016$ one-tail, on 30 subjects); in presentation 2, the effect was weakly confirmed ($P < 0{\cdot}08$ one-tail); and in presentation 3, the effect was not confirmed at all. In fact, in presentation 3, before-break clicks are responded to significantly *faster* than after-break clicks ($P < 0{\cdot}02$ two-tail). The difference between the presentation orders is primarily due to a decrease in reaction time to before-break clicks (71 msec. difference between the first and third presentations in Experiment I, 68 msec, in

Experiment II). Reaction time to after-break clicks is considerably less affected by practice (37 msec. difference for Experiment I, 38 msec. for Experiment II) ($P < 0\cdot02$ two-tail across subjects for the relative susceptibility of before-break clicks to familiarity in Experiment II).

The lack of prosodic cues in this experiment led to a higher error rate on correctly recalling the sentence, so each response was categorized according to whether the recall was substantially correct or incorrect. In presentation 1, 34 per cent of the sentences were not recalled correctly so it might be the case that the relative slowness of RT to before-break clicks is arbitrarily related to perceptual processing of the syntactic structure. To check for this, we tested this effect separately on those responses for which the sentence was correctly recalled and on those for which the sentence was incorrectly recalled. The relative slowness of RT to before-break clicks was confirmed across both sets of responses ($P < 0\cdot03$, one-tail for correct-recall responses; $P < 0\cdot02$ one-tail for incorrect-recall responses).

As in Experiment I, before-break clicks tended objectively to precede after-break clicks by one word. If a systematic decrease in reaction time were correlated with the number of syllables of the sentence that had occurred before the click then the tendency for before-break clicks to receive slow RT's might not be due to the syntactic structure. In Experiment II, there were enough subjects to test this effect for those sentence-syllable positions only in which both before-break and after-break clicks occurred (albeit in different sentences). When only overlapping sentence-syllable positions are included, the effect is confirmed, although less strongly ($P < 0\cdot06$, one-tail).

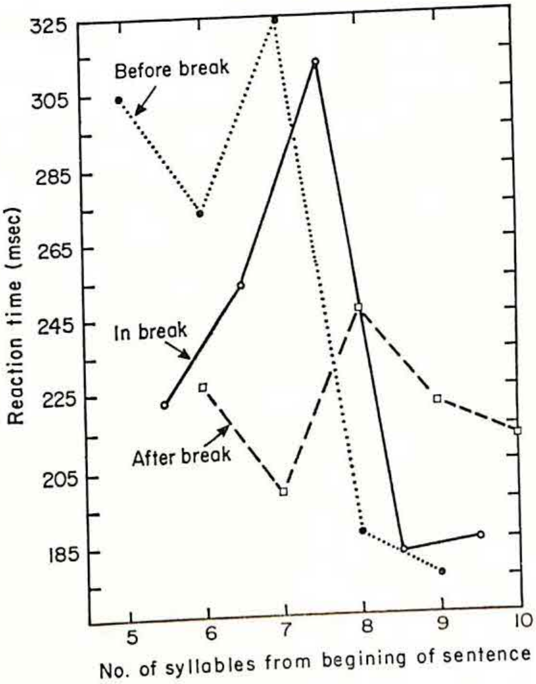


FIGURE 1. The effect of absolute length of preceding sequence on reaction time to clicks in the three different syntactic positions.

However, there was also a general interaction of the number of syllables in the sentence prior to the click with the relative slowness of RT to before-break clicks. This interaction is presented in Figure 1 which displays the RT's according to the number of syllables from the beginning of the sentence in which they occur. (For example, in sentence (1), a before-break click occurred in the fifth syllable, the in-break click between the fifth and sixth syllables and the after-break click in the sixth syllable.) Before-break response times are slower than after-break response times for clicks objectively located in the fifth, sixth or seventh syllables from the beginning of the sentence. That is, RT to before-break clicks are much faster than to after-break clicks for the fifth, sixth and seventh syllable positions ($P < 0{\cdot}001$ two-tail grouping the three syllable positions by subject). In contrast, before-break response times are *faster* than after-break response times for clicks objectively located in syllables eight, nine and ten: ($P < 0{\cdot}05$ two-tail by subject). The shift in the relative slowness of RT to before-break clicks appears to be due entirely to an increase in RT to before-break clicks in the later syllable positions (the difference between RT to clicks in syllable positions 5, 6, 7 and in 8, 9, 10 for before-break clicks is significant, $P < 0{\cdot}01$ two-tail by subject); the reaction time to after-break clicks is roughly constant for all syllable positions. The responses to in-break clicks also appear sensitive to the number of syllables from the beginning of the sentence that the click occurs. In-break clicks are responded to significantly faster than after-break clicks, for the clicks in the eighth, ninth and tenth syllable ($P < 0{\cdot}02$ one-tail by subject).

Several results indicate that the psychological mechanism for reacting to clicks is distinct from the mechanism for locating clicks. The pattern of errors in the subjective location of the clicks replicates the results of previous experiments. Seventy-two per cent of click location errors of ½ syllable are toward the major break in presentation 1 and 70 per cent in presentations 2 and 3. However, there is no difference in the confirmation of the click location hypothesis for those clicks in the first five to seven syllables and in the eight to ten syllables in presentation 1. That is, sentence-click combinations which evoke obviously distinct attentional strategies (as revealed by the different effects of syllable positions in presentation 1 and by the different effects of syntax on RT in presentations 1 and 3) confirm the effect of syntax on subjective click location to the same extent.

If the displacement of the reported location of clicks is due to immediate attentional mechanisms, then it might be expected that clicks responded to fastest should either have the least error in reported location or would systematically be reported as preceding clicks which are responded to slowly. The first possibility was tested by correlating separately the accuracy of each subject's location responses with his relative speed on those responses. There was no tendency for positive or negative correlation across subjects for responses in presentations 1 and 2 and a weak positive correlation in presentation 3. In presentation 3, 18 subjects showed a positive correlation, and 12 subjects a negative one ($P < 0{\cdot}07$ by Wilcoxon one-tail on the 30 correlations). This indicates that in familiar sentences, there may be a positive correlation between speed of reaction time and accuracy in subjective location of a click, but that in unfamiliar sentences the mechanisms for reaction time and click location are dissociated. To test the second possibility, the subjective location responses were divided into those which preceded the objective click location, and those that followed the objective location. The reaction time to those clicks subjectively reported as preceding their actual location was not faster than for those reported as following their actual location.

Discussion

The main goal of these experiments was to substantiate the earlier claims that syntactic knowledge is actively utilized in the perception of speech. We found that reaction time to clicks before clause breaks is affected by clause length and by familiarity with the sentence more than the reaction time to clicks after clause breaks. This indicates that syntactic structure does systematically modulate attention during speech perception. The fact that these effects occur in response to artificially constructed stimuli (as in Experiment II) shows that they are not due to any superficial acoustic cues of syntactic structure. Rather, listeners must actively provide the structural analysis which guides their attention as they hear each sentence. The results also indicate that attention to non-speech activity is reduced at the ends of clauses rather than enhanced: before-break clicks receive reactions which actually occur during the clause break, and it is these reactions which are the longest overall; furthermore, reaction times to clicks objectively in the break are not fastest.

These results are not consistent with the view that the points between clauses in a speech "channel" demand the least attention away from a click on a competing click "channel." Several recent experiments support a different view of the interaction of attentional processes (Kahneman, 1967. A similar view is proposed by Wundt in *Physiologischen Studien*, pp. 256 ff.). According to this view, carrying out the perceptual analysis of a stimulus is an *activity* which draws attention away from all other activities. Thus, on this view, competition for immediate attention is not between one sensory information input "channel" and another, but between *all* external sensory inputs and all internal perceptual analysis (see Bever 1968).

In sentences, the clause is a natural unit for internal perceptual analysis: during clauses we listen to the speech *and* non-speech stimuli: at the ends of clauses we encode perceptually what we have just heard. Accordingly, a click at the end of a clause is responded to relatively slowly because it coincides with the point of internal perceptual analysis of the preceding sequence. A click at the beginning of a clause is responded to relatively quickly because it conflicts with relatively little internal perceptual processing.

In both Experiments I and II the relation of the syntax surrounding a click and reaction time reversed for sentences as they became increasingly familiar. The reversal in the effect of syntax on reaction time in familiar sentences reflects a change in the point at which the internal analysis of the clause takes place. In the perception of a novel sentence, the listener organizes the immediately preceding stimulus material near the end of each clause. On the other hand, in the recognition of an already-learned sentence, its internal organization is already present in memory, and the listener's only task is to check which of the previously heard sentences it is. Recent work suggests that the beginning of a constituent predicts the end of that constituent better than vice versa (Johnson, 1966). Thus, a reasonable recognition strategy would be to match internally the beginning of each clause to the set of previously heard clauses to see which of those clauses it is: after comparing the beginning of the clause with the stored representation, the rest of the clause would be relatively predictable. Accordingly, familiarity with a sentence decreases RT to clicks at the end of a clause relatively more than familiarity decreases RT to clicks at the beginning of a clause.

Our interpretation of the effect of internal perceptual analysis on all external attention is also consistent with the fact that clause length decreases RT to end-of-clause clicks but does not affect RT to clicks at beginning of clauses in unfamiliar sentences. The predictability of the end of a clause increases with clause length. Thus, the listener can start his internal perceptual analysis of a long clause before the end of the clause. A click at the end of long clauses would thus occur just after the internal analysis is completed and while the listener is checking the prediction: thus clicks in these positions are responded to relatively fast.

The errors in click location are "attracted" to the break for both unfamiliar and familiar sentences and for sentences with short and long first clauses. Thus the segmental "integrity" of a clause is reflected both in the perception of new sentences and in the recognition of old ones, even though the initial perceptual mechanisms clearly differ. Previous interpretations of the tendency for clicks to be located between clauses (Fodor and Bever, 1965; have followed the proposal of an "n-channel" or "filter-amplitude" theory of auditory attention (Treisman, 1964). On this view, listeners pay attention to the speech except between perceptual units when the attention shifts to the click "channel." The effectiveness of the clause as a unit was demonstrated by the tendency to locate clicks between them allegedly as a function of the ability to switch sensory attention to the click "channel" only between such units.

However, the experiments described in this paper do not support the claim that sensory attention to non-speech stimuli is highest between clauses. This forces us to reconsider the nature of the previous results on subjective click location. If it is not the case that non-speech stimuli receive the highest sensory attention at the points between clauses, why do errors in location of clicks tend to migrate towards the clause break, and why is there a marked tendency for clicks objectively in the clause break to be correctly located?

Many of our results indicate that the attentional system tapped by the reaction time measure is distinct from the behavioural process which produces the systematic errors in click placement. Thus, the systematic nature of errors in click location are due to perceptual processing which follows the initial segmentation of the speech stimulus. We have argued that immediate reaction time interacts with the process of *developing* the internal perceptual organization of the speech. The systematic attraction of click location by clause breaks indicates that the "channel" view of perception is true for speech and clicks once the speech channel has received such initial perceptual organization. That is, listeners first organize the speech into major segments, then they relate the speech and click temporally: it is the latter process that maintains the integrity of the speech units as revealed in the location of clicks.

Certain features of these experiments require investigation with different techniques and study of additional theoretical variables. Reacting to a click in a sentence is itself highly disruptive and may introduce special perceptual properties. Also, the reversal of the relative slowness of RT to before-break clicks for sentences with relatively late clause breaks and for all sentences when they are familiar, indicates that factors such as sentence length and sequence predictability may be as relevant as well as syntactic structure. Nevertheless, these results do demonstrate that syntactic knowledge continuously modulates immediate attention during listening, as well as providing the basis for the subsequent perceptual segmentation of speech into discrete major units.

This work was supported in part by NDEA, AF 19 (628)—5705 to MIT and the U.S. Department of Defence, Advanced Research Projects Agency, Contract No. SD-187, to Harvard University, Center for Cognitive Studies, and in part by the National Institute of Mental Health, Grant No. Pol-MH12623 to Harvard University, Center for Cognitive Studies, by the Harvard Society of Fellows and National Institute of Child Health and Human Development, Grant No. HD 01368–03 to Klaus F. Riegel. The authors are particularly indebted to Dr. Merrill Garrett for his thought and advice. The authors are also indebted to Drs. D. Foss, D. Kahneman, C. Osgood, and P. Carey for reading the manuscript, to Henry Koopmans and Roberta Banay for assisting in the data analysis.

References

Abrams, K. (1965). Attention and Speech Perception: AB thesis M.I.T.

Bever, T. (1968). A survey of some recent work in psycholinguistics, in *Specification and Utilization of a Transformational Grammar*. Final report of Contract No. AF 19 (628)—5127 awarded to W. Plath, IBM, T. J. Watson Research Center, P.O. Box 218, Yorktown Heights, N.Y.

Bever, T., Lackner, J. and Stolz, W. (1969). Transitional probability is not a general mechanism for the segmentation of speech. *J. exp. Psychol.* **79**, 387–94.

Bever, T., Lackner, J. and Kirk, R. (1969). The underlying structure sentence is the primary unit of immediate speech processing. *Percept. Psychophysics* **5**, 225–34.

Fodor, J. and Bever, T. (1965). The psychological reality of linguistic segments. *J. verb. Learn. verb. Behav.* **4**, 414–20.

Garrett, M., Bever, T. and Fodor, J. (1966). The active use of grammar in speech perception. *Percept. Psychophysics* **1**, 30–2.

Garrett, M. and Bever, T. (1969). The perceptual segmentation of sentences. In Bever, T. and Weksel, W. (Eds.), *The Structure and Psychology of Language*. New York: Holt, Rinehart and Winston (in press).

Johnson, N. (1965). The psychological reality of phrase structure rules. *J. verb. Learn. verb. Behav.* **4**, 469–75.

Kahneman, D. and Beatty, J. (1966). Pupil diameter and load on memory. *Science, N.Y.* **154**, 1583–5.

Ladefoged, P. and Broadbent, D. E. (1960). Perception of sequence in auditory events. *Q. Jl exp. Psychol.* **12**, 162–70.

Treisman, A. (1964). Selective attention in man. *Br. Med. Bull.* **20**, 12–6.

Wundt, W. (1929). *Physiologischen Studien*, Vol. 2. Leipzig: H. Engelmann.

Received 23 *May* 1969

APPENDIX

The Sentences used in Experiments I, II and III

In addition to his wives, the prince brought the court's only dwarf.
They fought tooth and nail, to get past the huge angry crowd.
By making his plan known, Jim brought out the objections of everybody.
Since she was free that day, her friends asked her to come.
If you did call up Bill, I thank you for your trouble.
We asked the mean old man, to be kind to his dog.
That the matter was dealt with fast, was a surprise to Harry.
After the dry summer of that year, most crops were completely lost.
Any student who is bright but young, would not have seen it.

Q. Jl exp. Psychol. (1969) **21**, 291–298

BOOK REVIEWS

Animal Discrimination Learning. Edited by R. M. Gilbert and N. S. Sutherland. London: Academic Press. 1969. Pp. xvi + 501. 135s.

The papers in this volume grew out of the proceedings of a symposium held by the Experimental Analysis of Behaviour Group at the University of Sussex in 1967, some of them being solicited after the conference "in order to round off the discussion of certain topics." Much of the material, as one may expect, concerns the role of selective attention in discrimination learning. Chapters by D. R. Thomas, Honig, A. R. Wagner, J. M. Warren and B. McGonigle, and S. Siegel are written from this point of view, and contain discussion of such matters as the Overtraining Reversal effect, extra-dimensional discrimination training, and incidental stimuli. On the whole, the kind of theory associated with the work of Lawrence, and more recently Sutherland and Mackintosh, suffers rather badly at the hands of these investigators; but there is little doubt (as Thomas points out) that it had lead to a great deal of important work being done which would have been unlikely to grow from a less inspired theoretical basis.

Briefly, the other chapters contain discussions of: generalization of conditioned suppression (Hoffman), comparative studies of reversal and probability learning (an entertaining controversy between Mackintosh and Bitterman), Contrast and the peak shift (Bloomfield), PRE and frustration theory (a mathematical approach by Gray and Smith), Temporal discrimination (chapters by Harzem and Boakes), Mass action and split-brain learning (Russell), and Discrimination under avoidance schedules (Hurwitz and Dillow). Keehn and Gilbert contribute general discussions, and Sutherland advances his latest views on pattern recognition.

Most of the chapters are highly instructive, and the book as a whole is well produced with the exception of a few confusing errors. Figure 1 in Honig's chapter surely has its groups wrongly labelled, and the only way to make sense out of the second paragraph in Warren and McGonigle's discussion of the additivity of cues is to read "noncontinuity" for "continuity." In discussing an experiment by Wall, Harzem claims (p. 322) that "... the longer latencies, at 15 and 30 seconds after reinforcement, gradually increased," but Wall's data seem to show the exact reverse.

These minor points aside, a rather general criticism may be levelled against the treatment of Generalization. The central question is: how do we measure the sharpness of a generalization gradient? Relative sharpness of gradients is a vital theme in several Chapters, and we may take as an example Honig's claim that gradients are sharpened by true relative to pseudo-discrimination procedures. Of course, it is simple to show that two gradients differ in sharpness when the same number of responses is distributed between them, for this will involve the unambiguous criterion of the curves crossing, and thus differing on an ordinal scale. But if the putatively sharper gradient encompasses a larger number of responses this crossing may not happen, and then obvious difficulties arise. Honig adopts the common solution of assuming that preference is validly measured on an equal-ratio scale: that is, equal A/B ratios indicate equal preferences for A and B even when A + B is a variable. Thus: "Since slope differences may in part be due to different levels of response output in different groups, the generalization gradient of each subject was recalculated in terms of percentages." But what exactly is the theoretical or empirical justication for this step, which involves rejecting the possibility that response preference is reflected in *absolute* differences in output: that is, on an interval scale? Suppose we find, for example, that a particular stimulus presented in extinction reduces total responding on two levers, but increases the percentage responding on the lever producing the stimulus while leaving *absolute* differences in responding on the two levers constant. Should we conclude that the stimulus reduces response output and is sumultaneously preferred, or that it has no effect on preference as measured by an interval scale? The choice of scale clearly determines the interpretation of the data. This describes, by the way, the result of one of my own (unpublished) experiments, so I am not raising the issue just for fun.

Actually, it must be said that some of Honig's gradients cross even in their unadulterated form, so there may be no particular difficulty with his data. Where genuine problems arise is in the chapter by Hoffman, who claims that generalization gradients of conditioned suppression are sharpened in the course of extinction. These gradients involve differences in absolute level of responding (inevitably so, since extinction of suppression is proceeding), and there is only one case of cross-over in a series of 35 points defining 5 gradients. Nevertheless, Hoffman appears to argue that the later gradients are sharper than those measured in extinction. Certainly, these later gradients *look* sharper, but there is no reason why this should not be a consequence of the increased number of responses to be distributed. Hoffman himself seems to arrive at a similar conclusion when he derives his results from the assumption that the rate of recovery of responding after suppression is positively accelerated: but it is hard to reconcile this with his more general claim to have proved that generalization is not a failure to discriminate. The results really prove nothing about discrimination either way: only that the shape of a gradient can be affected by the absolute level of responding.

The problem of detecting interaction in the absence of crossover has been discussed in a paper on conjoint measurement by Luce and Tukey (1964)† and I am grateful to S. E. G. Lea for drawing this paper to my attention. Mr. Lea further points out that a preliminary application of Luce and Tukey's measurement procedure to Hoffman's data fails to show any significant interaction between generalization and extinction. This supports the view that the changes in "shape" of the gradients during extinction are perceptual rather than mathematical.

M. J. MORGAN

---

† Luce, R. D. and Tukey, J. W. (1964). Simultaneous conjoint measurement: a new type of fundamental measurement. *J. math. Psychol.* 1, 1–27.

Discrimination Learning. By D. A. Riley. Boston: Allyn and Bacon Inc. 1968. Pp. 172. Paperback 21s.

This book is part of a new series written with the object of surveying selected topics in depth: the series is aimed primarily at undergraduates, but the present book will be of interest to research workers too. The author reviews certain aspects of three problems—stimulus generalization, transposition and the continuity non-continuity dispute. He emphasizes throughout the relationship between experiment and theory, and examines in careful detail whether each experimental result quoted really can be taken as supporting this or that theoretical position. He sets out the details of various experiments much more fully than is common in textbooks, and he writes clearly about the historical development of the different problems. As far as it goes, this is a good book and should be ready by anyone interested in discrimination learning, but despite its virtues, it could have been better. It may be worth examining some of its shortcomings in more detail since whereas its virtues are its own, its defects are common to most books in the field of animal psychology.

First, the author does not appear to have any firmly held convictions of his own, and as a result, the book makes rather flat reading. He repeatedly cites four or five theories, shows that it is very difficult to be sure whether this or that experiment supports this or that particular theory and concludes that we do not have enough evidence to decide between the theories. All theories and all experiments are accorded equal status and treated with equal solemnity; the author gives no emphasis to what he thinks is good or to what he thinks is bad.

Although extensive detail is often useful, it can prevent the reader from seeing the wood for the trees. Both Lashley and Wade's (1966) experiment on generalization and Lashley's

(1942) experiment on irrelevant cue learning are treated in great detail. Yet both experiments are only of historical interest and were so badly designed as to be almost worthless; moreover, the issues have in any case been clearly resolved by later experiments. Some of the most crucial of these are not even mentioned. In view of Terrace's (1963) results, it is really quite pointless to argue about whether an animal can learn about one cue on trials when its behaviour is controlled by another. Yet these results are not cited in this connection, though they are mentioned elsewhere in the book.

Riley's evaluation of theories is rather curious, since each theory is (deliberately) treated only in relation to the narrow range of experiments actually discussed in the book. Theories are never examined to see what range of facts they encompass, nor is their plausibility or elegance considered. The reader is never made aware that the subject deals with organisms having an evolutionary history and that the way in which they work has been determined to a considerable extent by evolutionary pressure.

The discussion is not always as clear as it might be. The author has a tendency to take over problems that have been raised by others and to discuss them at length without considering whether they are really worth discussing. For example, the section on generalization revolves around the question of whether generalization is simply failure to discriminate or whether it represents response to discriminated similarities. Riley (in common with previous psychologists who have raised this issue) fails to clarify the issue. He writes "one view assumes that increasingly great differences become more detectable; the other that increasingly great differences, all of which are detected, can be scaled for magnitude of difference." One can understand what the distinction represents in man—if a generalization gradient is obtained from a human subject, then presumably the difference between the stimuli is "detected" if the subject is able to verbalize it. What it means for an animal to detect a difference to which it does not respond is not clear.

Perhaps some of the book's weaknesses stem from the author's avowed predilection (pp. 13–14) for explanation in terms of principles rather than in terms of mechanisms. He contrasts physiological explanations of behaviour with explanations (like Hull's) in terms of general principles, without apparently realizing that there is a third way of explaining behaviour, namely, by specifying the nature of the information-processing mechanisms at work in the organism, without making physiological identifications. It is curious how little modern developments in our understanding of information processing have affected the way psychologists think about animal learning. This book, in common with many others, gives the impression that psychologists studying animal learning spend their time taking in one anothers' washing rather than trying to discover how behaviour is controlled.

Despite these critical remarks, this is a good book as books on animal learning go and should be read by everyone interested in the field.

N. S. SUTHERLAND

Annual Review of Psychology. Vol. 20. Palo Alto: Annual Reviews Inc. 1969. Pp. ix + 544. $9.00.

"From the Middle East, following in the tradition of Abraham (4), Moses (136) and Solomon (229), four interesting books have been published [Matthew (129), Mark (126), Luke (118), John (97)]: these are thorough developmental studies (birth to death!) of a single subject, and support the so-called "Trinity" theory of Christianity, while confirming a number of predictions by Ezekiel (50), Amos (7) and Jehovah (personal communication). However they overlap considerably in the material they present, the adolescent period is inadequately covered, and no significance levels are given for the miracles. . . ."

The style of the *Annual Review of Psychology*, with its mass of references, obscurities and superficial criticisms is all too easy to parody, and therefore it is a pleasure to welcome the current volume, where so many of the chapters are easy and informative to read. This year's topics are Developmental Psychology, Basic Drives, The Physiological Basis of Memory, Audition, The Chemical Senses, Perception, Colour Vision, Personality, Attitudes and Opinions, Electrical Stimulation of the Brain in Behavioral Context, Neuropsychological

Bases of Schizophrenia, Human Abilities, Instructional Psychology, Personnel Selection and Scaling.

Foremost among the successful chapters are Flavell and Hill's Developmental Psychology, which has some illuminating comments to make on the differences between American and Genevan research, and Zinnes' amusing and provocative chapter on Scaling, which looks forward to the extinction of such phrases as "fallible data" and "the psychophysical law."

Of the less successful chapters, von Fieandt and Wertheimer's review of perception tries to cover too much ground, and in places reads like an itinerary for a tour of the world's four hundred most beautiful psychology laboratories. I found Ripps and Weale's chapter on colour vision too technical, but no doubt this is at least as much my fault as theirs.

My impression is that the range of an individual volume of the Annual Review gets narrower every year, not only in the topics selected, but in the individual author's treatment of their topics, and while this seems the only course of action if chapters are to remain intelligible (seven plus or minus two references per page is the average of the most successful chapters in the present volume), it does mean that a comprehensive annual survey of psychology within the confines on a single volume is no longer possible. Those of us whose channel capacity for journal reading is small will regret this.

An engaging feature of many of the chapters is the profusion of ingenious apologies authors make in their introductions to those whose work is not included in their references, and the present reviewer, in omitting reference to other chapters in this volume can only echo Green and Hemming's elegant apology at the beginning of their chapter: "The authors that are slighted may be consoled by the observation that truly original contributions are seldom quickly appreciated."

Philip T. Smith

Cutaneous Sensation. By David Sinclair. London: Oxford University Press. 1967. Pp. xii + 306. 63s.

The study of cutaneous sensation is of prime importance to anyone who is concerned with the central nervous system. A knowledge of its mechanisms can illuminate the diagnosis of neurological disorders and is basic to an understanding of pain. Such is the view expressed by Professor Sinclair in this excellent book in which he collates a vast amount of information from the many disciplines concerned with this subject. Drawing upon anatomical, physiological, psychological and clinical findings with regard to cutaneous sensation, he provides a reference work which is probably unique and which will certainly be invaluable to all those engaged in sensory research.

The book is divided into four main sections. Section I (General Background) covers the three major theories of cutaneous sensation and anatomical, physiological, psychological and mathematical approaches to its investigation. Section II (The Apparatus) gives a detailed account of the anatomy of nerve endings, the first sensory neurons and the central pathways and contains some excellent illustrations. Section III (The Sensations) deals with the sensory correlates of the peripheral and central neural apparatus, introspection and sensory quality, stimulus and sensation, and spatial and temporal factors. In Section IV (General Review), Professor Sinclair critically evaluates the "specificity" theory, the "duality" theory and the "pattern" theory of cutaneous sensation in the light of the experimental evidence and comes to the conclusion that the modified pattern theory of Melzack and Wall holds the greatest promise for further productive research.

Professor Sinclair is of the opinion that current thought concerning the sensory nervous system is unduly dominated by the techniques of electrophysiology which many people seem to regard as more "scientific" than other methods of investigation. He points out that the findings of the electrophysiologist, no less than those of research workers in other disciplines, are very much at the mercy of particular experimental conditions, such as the type of electrode, the nature of the anaesthetic and the kind of stimulus used.

Throughout the book the reader's attention is drawn to the many pitfalls which beset the experimenter's path and there are some very salutary criticisms of methods and techniques. Research workers are urged not to put all their eggs into one basket—electrophysiology—

and, since sensation is by definition a subjective problem, Professor Sinclair makes a plea for the use of introspection. He believes that a revival of enthusiasm for some existing methods is needed rather than the invention of new ones. His own clear presentation of the problems and his synthesis of many small areas of study, should go a long way towards generating this enthusiasm.

Two shortcomings of this otherwise splendid work are that there is no author index and the large reference section does not give the relevant page numbers.

DIANA R. HASLAM

Modelling of Thinking and the Mind. By N. M. Amosov. Translated by L. Finegold. New York: Spartan Books. London: Macmillan. 1967. Pp. xiii + 192.

The philosophical standpoint and the general assumptions of modelling living organisms at different levels of complexity are the most interesting features of this book. The total activity of the most complicated systems such as human brain is "no more than a summation of activities in its simple elements." Hence the author's optimism in modelling and creating of "machines which have programs equivalent to that of the brain" (p. x). Contemporary science (even psychology) must "depart from qualitative descriptions and proceed to quantitative models" (p. xiii). Amosov's general hypothesis of modelling is based on algorithms. Its principle is hierarchical information processing, hierarchical program of activities and hierarchical structure of systems.

Only two of the five chapters deal with the modelling of mind. The book is an attempt to explain the development and mental activity of living organisms in informational terms without paying any attention to anatomy and physiology. The processing of information is studied at the molecular, unicellular and multicellular level. The bases of any higher cortical activity are the "programs of attention": some pieces of information are "amplified" (Russian: *usileny*, i.e. intensified) others suppressed; the relationship to Pavlov's and Uchtomski's theories about "domination" is obvious. The suggested descriptive models of feelings, sensations, creative activities, thinking and other brain functions are clearly the conception of a clever cyberneticist but a superficial psychologist. We have great difficulties in defining the concept of thinking; according to Amosov, however, "the concept of 'thinking' should not present any great difficulty nor should it require extensive explanation" (p. 140). Amosov does not avoid the fault (so common to many of his compatriots) of using loose terms with no specific meaning applied arbitrarily to various diverging concepts.

Both the original work and its translation are without bibliographical references which makes it even more difficult to assess Amosov's original contribution.

The painstakingly literal translation of Mr. Finegold does not cover up the inveterate Soviet habit of trying to express materialist notions in idealistic language.

IVANA MARKOVÁ

The Mind of Mnemonist. By A. R. Luria. London: Jonathan Cape. 1969. Pp. xi + 160. 25s.

This short but exquisite account of one man studied over a period of nearly thirty years may prove to be one of Professor Luria's most important and influential contributions to psychology. Here, for almost the first time, it is possible to watch the process of remembering from within, to watch it at work in slow motion and to analyse the changes that occur with maturation; for the subject of the study "S," carefully encouraged and drawn out by the observer, is conscious of and able to describe processes which probably occur to some extent in everyone at a pre-verbal age and level, and which are therefore particularly difficult to analyse.

The processes which S used to retain his material—imagery in all modalities—is typical of what is believed to happen in the young child; it is interesting that when S wanted to extend his range, he had to contrive for himself a way of concentrating groups of events into symbols and eventually blotting out some by "forgetting" them—processes which are usually carried out by most people at an unconscious level.

The limitations of imagery for logical thinking and the difficulties that arise in dealing with sequential arguments if the incoming information is decoded in small chunks are beautifully demonstrated, but despite—or perhaps because of—the information contained in this book the nature and function of memorizing remain as obscure as ever. S's memory processes did not obey any of the "laws" which we have grown to regard as sacrosanct. They did not depend on rehearsal, they were not subject to retro or pro-active inhibition. The passage of time did not even dim them. There was no effect of serial position. If we accept that mental processes can be envisaged as something like the output of computers, S's computer had unlimited storage space.

Unfortunately we learn even less about forgetting than we do about retention. When S did finally find a means of erasing images (not so much, it seems, to gain extra storage space, but to avoid confusion during successive stage performances) the process remains as the author says "as unfathomable to him as it did to those of us who were studying him." S merely decided to "will" an image away. "If I don't want a chart to show up it won't." And it didn't.

Power to form and erase mental images which had the same vividness as reality was not the only part of himself over which S had voluntary control. He could lower or raise his body temperature, alter the speed of his heart-beat, and blot out pain as desired. Perhaps it is not surprising—and certainly the present reviewer would not consider it as a slur on his personality or an indication of his being a failure—that he opted out of the rat-race for notoriety and prestige and decided merely to live life as it came to him. MOYRA WILLIAMS

Determinants of Infant Behaviour IV. Edited by B. M. Foss. London: Methuen. 1969. Pp. 303. £3 10s.

A most noticeable feature of developmental psychology during the last ten years or so has been the rapid increase in interest in human infant behaviour and also the atttempt to link data on this behaviour to equivalent work on sub-human species. The series of seminars on Mother-Child interaction held between 1961 and 1965 under the auspices of the Tavistock Institute accurately reflects these twin trends and indeed has done a great deal to foster them. The books produced by these seminars have made a significant contribution to progress in this area, as much by generating enthusiasm as by spreading information.

This is the fourth and final book resulting from these seminars. Appropriately it is divided into two sections, one on animal, the other on human infant and maternal behaviour. Although each section is relatively homogeneous there seems to have been little attempt to make a link between the two sections. This is a pity since the purpose of these seminars was to make this kind of contact. Perhaps the contact was made in discussions of these papers but this is not apparent since in this book, in contrast to previous volumes, very little of the discussion is given, and that only on one of the papers.

The animal section contains three papers on primate infant mother interactions and two on the reaction of virgin and non-virgin rodents. The first primate paper, that by Hamburg reports behaviour in the field, and is extremely difficult to evaluate. The author makes frequent generalizations without detailed evidence and refers to evidence (does everybody besides myself know what the Reynolds' study was?) without giving references. The other two primate papers, by the Harlows and by Hinde, are admirable. They both contain clear accounts of continuations of lines of research that are already deservedly well known. The two rodent papers are remarkably similar in that they both examine the reaction of virgin and non-virgin females to infants. The difference between the two papers is in their main dependent variable, which in Lott and Rosenblatt's papers is retrieval latency, while in Richards' paper, somewhat bizarrely, it is the number of infants killed and eaten by these females.

The human section is somewhat less satisfactory for a final report. Two of the papers, those by Wolff and by the Gewirtzes are plainly preliminary reports. One other by Decarie on thalidomide children is very specialized. Another paper by Sander must be dismissed as being altogether too pretentious. On the basis of nebulous evidence he puts forward an abstract and meaningless theory that the two most important factors in development are adaptation and information processing. Another contribution, by Morgan and Ricciuti, is interesting but confounds variables: this experiment concerns the effects of strangers on infant behaviour. The two experimenters (one a man, the other a woman) took it in turn to be the stranger and observer. The two experimenters had significantly different effects on the infants: this is attributed to their effect as strangers and the alternative explanation, that the difference was in the way that the two experimenters observed, is not considered.

Altogether this is a patchy but interesting report. What is needed now is a book by a single author in which the contribution of all four seminars are brought together in a single framework.

P. E. BRYANT

Brain and Behaviour. Volumes 1–4. Edited by Karl Pribram. Harmondsworth: Penguin Books Ltd. 1969. Pp. 496 + 573 + 523 + 524. 12s. 6d. per volume.

This is a curious selection of readings with some notable virtues. R. S. Peters and Pribram speak on the subject of Freud while Freud speaks on Aphasia. I regard this as an excellent division of labour. The series is organized into four volumes: Mood, States and Mind, Perception and Action, Memory Mechanisms and Adaptation. In only the third is the title a guide to the contents, and the subsections are similarly unconstrained. About one-third of the papers are authored or co-authored by Pribram himself. Some of these serve the function of summarizing and linking the sections rather better than the usual comments of editors. Others, however, do not appear to have the generality or far-reaching significance usually required of reprinted papers. Along with modern work in neuropsychology are classical papers by Gall and Spurzheim, Coghill, von Monakow, Herrick, and others. On balance this is a useful selection.

MARK P. HAGGARD

Publications Received

Books

A Psychological Approach to Abnormal Behaviour. By L. P. Ullmann and L. Krasner. Englewood-Cliffs, N.J.: Prentice-Hall. 1969. Pp. xiii + 687.

A Handbook of MMPI Group Profiles. By R. I. Lanyon. London: Oxford University Press. Minneapolis: Minnesota University Press. 1969. Pp. vii + 78. 72s.

Problems of Psychology. By K. Dallett. New York and London: John Wiley & Sons Inc. 1969. Pp. vii + 349. 53s.

Rorschach with Children. By J. Francis-Williams. Oxford: Pergamon Press Ltd. 1969. Pp. xi + 167. 30s.

Symbols in Society: By H. D. Duncan. London: Oxford University Press. 1969. Pp. xvi + 252. 57s. 6d.

The Role of Learning in Psychotherapy. Edited by R. Porter. London: J. & A. Churchill Ltd. 1968. Pp. xii + 340. 60s.

Paperbacks

Deviance: Studies in the Process of Stigmatization and Societal Reaction. By S. Dinitz *et al.* London: Oxford University Press. 1969. Pp. 575. 36s.

Multiple Choice and Retention. By H. Eklund. Uppsala: Svenska Bokforlaget. 1968. Pp. 223.

Children's Descriptions of Simple Events. By T. Nummenmaa. Helsinki: Suomalainen Tiedeakatemia, 1969. Pp. 61.

Q. Jl exp. Psychol. (1969) **21**, 299–311

THE PERCEPTION OF BRIEFLY EXPOSED POINT-SOURCES OF LIGHT

B. DELISLE BURNS, G. MANDL,† R. PRITCHARD‡ AND CORDELIA WEBB

National Institute for Medical Research, Mill Hill, London, N.W.7

Point-sources of light (dots) were exposed for 10 to 50 msec. before five dark-adapted subjects in a dimly illuminated room. During voluntary fixation with one eye, the target was exposed some 10° on the nasal side of the optic axis. The intensity × duration of all targets was 2 × threshold and they consisted of either a single dot, or a pair of dots separated by a distance that was less than that required for two-point discrimination. In two-thirds of trials both the single-dot and the two-dot targets were perceived as short thin lines of various orientation. Although individual percepts were unpredictable, there was a preferred or most likely orientation for responses to the single-dot target; this was near to the horizontal for all five subjects. There was no significant difference between the preferred orientations for single-dot targets tested at sites more than 1° apart in the visual field. When two single-dot targets, separated by about 1°, were exposed simultaneously, the orientations of the perceived lines sometimes differed by as much as 80°; occasionally, one target was reported as a dot while the other was seen as a thin line. If the single-dot was briefly exposed between two continuously visible and parallel straight lines, the target usually appeared as a thin line, parallel to the framing lines. Some of these results appear to be consistent with the hypothesis that the human visual cortex, like that of the cat and monkey, contains neurones that are orientation specific.

Introduction

Hubel and Wiesel first demonstrated the orientation preference of cells in the visual cortex of anaesthetized cats. Using extracellular microelectrodes, they recorded the discharges of cortical units, which would respond to brief exposure of a straight white line or black/white border in the visual field, provided that its orientation lay close to some optimal value. It appears that some degree of orientation preference is a property of most, if not all, such cortical units (Hubel and Wiesel, 1959, 1962; Burns and Pritchard, unpublished; Barlow, Blakemore and Pettigrew, 1967; Campbell, Cleland, Cooper and Enroth-Cugell, 1968). Hubel and Wiesel (1963) have also shown that visual units of similar orientation preference are arranged in columns, running radially through the cortical grey matter. They estimate the diameter of each column of functionally similar cells to be approximately 500 μ. There seems to be general agreement that the orientation preferences of cortical cells in the cat show no particular bias, and columns with every direction of preference are equally likely to be encountered. Similar orientation preferences have been demonstrated for cells in the visual cortex of the monkey (Hubel and Wiesel, 1968). Since the receptive fields of neurones in the lateral geniculate body are circular (Hubel and Wiesel, 1961), the behaviour of cortical units in the cat and monkey must be determined by local structural arrangements. In the rabbit this coding of sensory information seems to be performed earlier in the visual pathway; Barlow, Hill and Levick (1964) have shown that in the rabbit some of the retinal ganglion cells are orientation sensitive.

---

† Present address: Department of Physiology, McGill University, Montreal, P.Q., Canada.
‡ Present address: Department of Psychology, McMaster University, Hamilton, **Ontario**, Canada.

These observations on experimental animals suggest that neurones in the human visual cortex may also display orientation preference. In fact, there is already some indirect evidence to support this idea. Andrews (1967) has shown by psychophysical techniques that acuity for orientation and constant errors of perceived orientation of a short bright line are related to fixation point, orientation and exposure duration. His subject's task was always to set the short bright line parallel to a fixed line nearby. Under these optimal viewing conditions the true inclination of the line sometimes appeared to differ from that of the comparison line by as much as 30°, and deviations of 15° or so were common: acuity for orientation was optimal when the stimulus objects were horizontally or vertically orientated. In addition, the perceived shape of the short bright line was subject to illusory distortions on up to 20 per cent of the brief stimulus presentations; these distortions were as frequent but less extreme when the stimulus was continuously presented. Andrews interprets his results in terms of units in the human visual system that respond selectively to lines at particular orientations. He proposes that such units are most selective when "tuned" near the horizontal and vertical directions and are subject to long-term adaptation.

Gilinsky (1967) found that prolonged observation of a bright grating made of straight black and white bars caused a subsequent elevation of visual threshold for a test grating of the same orientation. This procedure produced a reduction in threshold for gratings subsequently exposed at right angles to the conditioning grating. It seems possible that in these experiments, the conditioning grating had selectively fatigued neurones with a particular orientation preference, thus raising their threshold to subsequent excitation by patterns of this orientation. The apparent facilitation of neurones possessing an orientation preference at right angles to that used during conditioning, is harder to explain.

Campbell and Kulikowski (1966) have made similar observations. Human subjects were asked to observe a grating of parallel, straight, light and dark lines; this grating was used to mask another simultaneously exposed grating that was less bright. The angle between the masking and test gratings could be varied. Campbell and Kulikowski measured the threshold for visibility of the test grating and found this threshold to be highest when both gratings were presented with the same orientation. As the angle between the test and masking gratings was increased, the masking effect fell off exponentially; at 12° on either side of a vertical test grating the masking effect was reduced by a factor of two. These results can be interpreted as the consequence of occlusion in cells of the human visual cerebral cortex, which possess orientation preference. Campbell and Kulikowski did not find the lowering of threshold at 90° to the conditioning grating that has been reported by Gilinsky.

We have described below, a different class of experiment, the results of which can also be interpreted as the consequences of exciting units in the human cerebral cortex that display orientation preference. The origin of these experiments came from the following somewhat tenuous reasoning:

if (1) cells in the human visual cortex display orientation preference, and

if (2) these units are arranged in functional columns of neighbouring units showing similar preference, and

if (3) a just perceptible point source of illumination in the visual field produces cortical excitation that is restricted to one column of cells, and

if (4) the human nervous system makes use of the information provided by the orientation preference of visual units,

then a point source of light exposed tachistoscopically at near threshold brilliance should appear as a thin white line.

Despite the uncertainties of the argument that led to these tests, our results have shown that a briefly exposed point of light was perceived by all five subjects as a thin white line in approximately two-thirds of exposures.

Methods

Figure 1, which is not drawn to scale, illustrates the equipment used in these experiments. The subject was seated comfortably with chin and forehead against firm, shaped rests. Subjects were asked to fixate a neon-illuminated target (F) on a pearl glass screen (G), 128 cm. in front of the left eye; both eyes were open but the right eye was prevented from seeing the glass screen by an opaque screen (S_1). The targets used for brief exposure in these experiments were images of the ends of three light-tubes projected through the lens system (L) upon the back-side of the pearl-glass screen. Only one of these light tubes (LT) is shown in the diagram. These tubes carried light from the spots of a Tektronix double-beam cathode ray tube (CRT) with P7 phosphor. The times for which the CRT spots appeared beneath the receiving end of the appropriate light-tubes were governed by Tektronix pulse and wave-form generators. In this way we were able to produce measured exposures of either a single dot of light (no. 2) or a pair of dots (nos. 1 and 3). The single dot was always projected midway between the other two and its intensity was arranged to be twice that of the pair. The separation and magnitude of the pair of dots was adjusted to be less than the limit for two-point discrimination in the part of the visual field used. Most of the work was carried out with targets centred 11° on the nasal side of the field on a horizontal line through the optic axis. In these experiments the dots measured 0·5 mm. diameter,

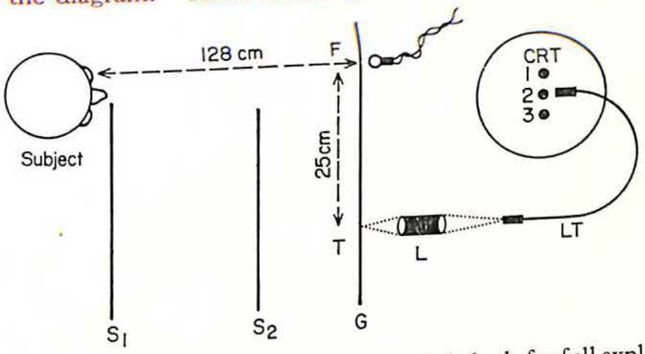


FIGURE 1. Diagram of the apparatus used. See Methods for full explanation.

with centres of the outer pair separated by 2·5 mm. The orientation of the pair of dots [Fig. 1(1 and 3)] exposed simultaneously, could be altered without knowledge of the observer.

Two out of the five subjects were naive volunteers, who knew nothing about the nature of the experiment or its objective at the time they were tested; they were nos. 2 and 5. The laboratory was dimly lit and all subjects were allowed 20 min. of dark adaptation before each experiment was begun. The illumination of the front of the pearl-glass screen by reflected light was approximately 0·01 milliLamberts. (If the subject normally wore spectacles these were used during the experiment.) The first step in every experiment was to estimate the subject's threshold for visibility of the targets that would be presented; this was measured as the threshold duration for a target of constant intensity. Exposure durations were then set at 2 to 3 times threshold for use during the experiment; the exposures used, lay between 10 and 50 msec. The instructions for a particular experiment were read to the subject and the experiment was begun. Test exposures were made approximately once in 5 sec. Before each exposure the operator said "Ready?", and roughly 1 sec. later he lifted screen S_2 and exposed the target. Subjects were asked to report the orientation of any line they perceived in terms of 'times of the clock', between 12-o'clock (vertical), through 3-o'clock (horizontal) to 6-o'clock (vertical). Thus a line appearing as 'vertical + 45°' would be correctly reported as 1·30. Most subjects were given a rest after each batch of 30 trials.

General

Rough tests have shown that results which are similar to all those described below can be obtained from targets exposed anywhere within the visual field. However, one requirement of these experiments was that the two-dot pattern should not be large enough to permit two-point discrimination and would therefore appear to subjects as a line. Since it proved much easier to work with relatively large patterns, we chose to perform all the main experiments in the peripheral field, some 10° on the nasal side of the optic axis. In fact, the phenomena described below do not appear to be critically dependent upon target position, intensity and duration of exposure, or size.

Experiment I: How often is a single dot perceived as a line?

The instructions to the subject ran as follows: "In each trial you will be presented with *either* a pair of dots in various orientations, which will appear to you as a uniform straight line, *or* one single dot. Immediately after each exposure, please report which of these two targets you saw, as a line or a dot. The targets will be presented to you in a random order, and we shall not necessarily present equal numbers of dots and lines."

In fact the numbers of dots and lines presented were equal, but they were exposed in a random sequence; the orientations of the 'lines' provided, also formed a random sequence. Most subjects received 120 tests including 60 dots and 60 lines, with orientations chosen at random from 1, 2, 3, 4, 5 and 6-o'clock.

Table I shows the results of this test:

The last row of the table shows the mean performance of the five subjects. While the line-figure (of two-dots) was perceived correctly in two-thirds of the trials, the single dots were perceived incorrectly as lines in about the same fraction of tests. The results show clearly that the subjects were unable to distinguish between the two targets.

During Experiment I, the subject was asked to report only whether he had seen a dot or a line; he was not asked to describe the orientation of any line that he perceived.

Experiment II: Is there a preferred orientation for the lines perceived when a dot is shown?

Experiment II, which is described fully below, was similar to Experiment I, except that the subject was asked to state the orientation of lines that he saw. Thus, there is some justification for comparing the results of Experiments I and II, and this has been done in the last column of Table I.

TABLE I

Subject			Lines			Dots			% of dots seen as lines		
Subject	Spectacles worn	Target exposed (msec.)	No. of lines shown	No. seen as lines	No. seen as dots	No. of dots shown	No. seen as dots	No. seen as lines	Experiment 1	Experiment 2	Experiment 1 and 2
1	No	50	60	50	10	60	11	49	82	75	78
2	No	12	28	13	15	32	16	16	50	75	62
3	Yes	40	60	31	29	60	28	32	53	47	50
4	Yes	50	60	42	18	60	25	35	58	65	62
5	No	40	60	43	17	60	21	39	65	68	66
Mean %					32				62	66	64

For Experiment II the instructions to the subjects were as follows: "You will be presented with *either* a very short line at a variety of angles *or* a single dot. Please report the angle of any line you see (as 1-o'clock, 2, 3, 4, 5 or 6-o'clock) or, if you see one, report a dot. The number of lines and dots presented may not be equal."

Figure 2 is constructed from the results obtained with subject no. 3. The targets presented to the subject are listed along the abscissa; the subject's report is plotted as ordinate. The entries in the last column show that this subject was most likely to perceive a dot as a line with orientation of 120 to 150° (4 or 5-o'clock).

If the subject had correctly reported all the targets shown, all the points in Figure 2 would have fallen in the shaded squares. His results for the line (two-dot) targets show some positive correlation between the orientation presented and the orientation perceived. Of the line targets he perceived as lines, 12 were reported with the correct orientation. If he had been uninfluenced by the orientation of the target, one would have expected a correct report on one-sixth of the 27 trials.

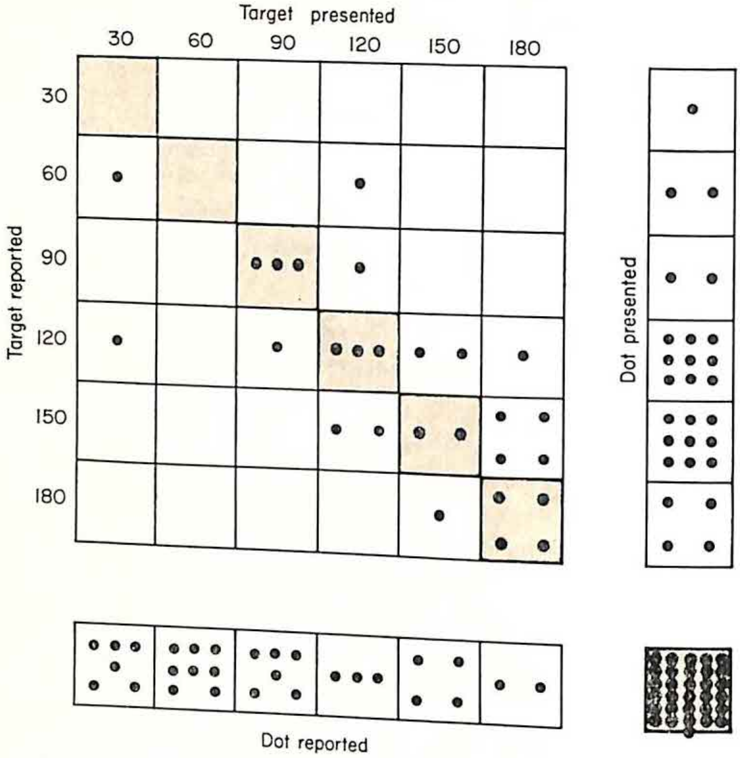


FIGURE 2. Results obtained from subject 3 in Experiment II.

The results of this experiment for all five subjects are shown in Table II.

A summary of the responses of all five subjects to presentation with single dots is provided in Figure 3. It will be seen that the orientation preferences were not the same, but most of them lay near to the horizontal.

TABLE II

Subject	Lines perceived	Correct reports of orientation	% Correct reports
1	52	17	32·7
2	25	6	24·0
3	30	12	40·0
4	44	9	20·5
5	53	12	22·7
Expectation if no correlation			16·5%

The thin white lines that were perceived appeared to have lengths ranging from 5′ to 30′ (2 to 12 mm. at the 128-cm. range of our screen) and rough tests showed that the greatest lengths were perceived when the orientation reported was close to the preferred orientation.

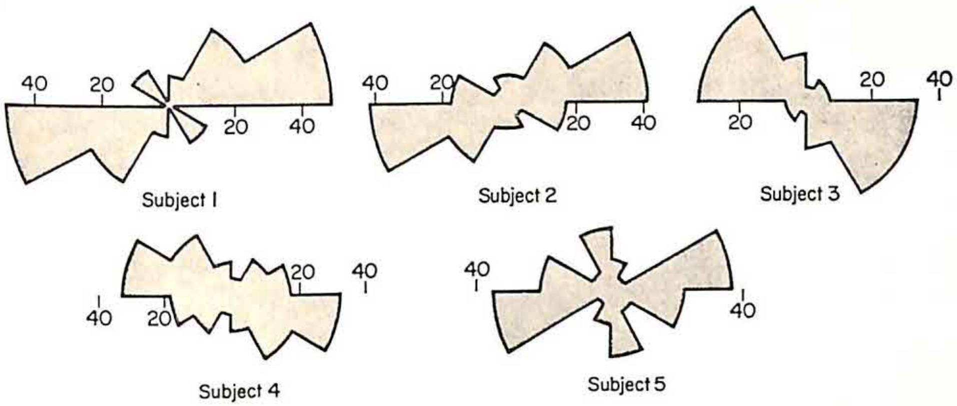


FIGURE 3. Summary of responses to dot-targets: Experiment II. Polar diagrams showing the orientations reported by 5 subjects observing dot-targets. Radial axis indicates percentage of all reports of a particular orientation.

Experiment III: Does the preferred orientation for the lines seen when dots are presented one at a time, vary with the position of the target in the visual field?

Hubel and Wiesel (1968) have shown that the preferred orientations of neighbouring neurones in the visual cortex of the monkey can be very different. It was therefore expected that the preferred orientations for responses to our targets would vary with varying positions in the visual field.

The instructions to the subject were the same as those given before Experiment II. In this experiment, however, the targets were presented at one of two neighbouring positions in the visual field, chosen in a random sequence. The subject was unaware that two positions for the targets were employed. The two positions used were 1° 15′ apart; both sites lay in the medial field on a horizontal line through the optic axis, about 10° from it.

In no case did the preferred orientations for the two sites tested differ significantly from one another.

Experiment IV: Are the preferred orientations for the lines seen, when two separate dots are presented simultaneously at two different sites in the visual field, always identical?

It appeared that the results of the preceding experiments might be the consequence of either astigmatic distortion of the retinal image, or of the small eye movements of physiological nystagmus. This and the subsequent experiment were designed to test these alternative explanations.

The instructions to the subject ran as follows: "In each trial you will be presented with two stimuli, simultaneously. Please report the apparent angle of both stimuli; the upper stimulus should be reported first." Two single dot targets were presented simultaneously; they straddled a horizontal line through the fixation point and were 10° on the nasal side of the fixation point. They were separated vertically by 1° 32′.

Results from subject no. 1 are shown in Figure 4. (This subject provided estimates of orientation that included 'half hours'.) If the upper and lower

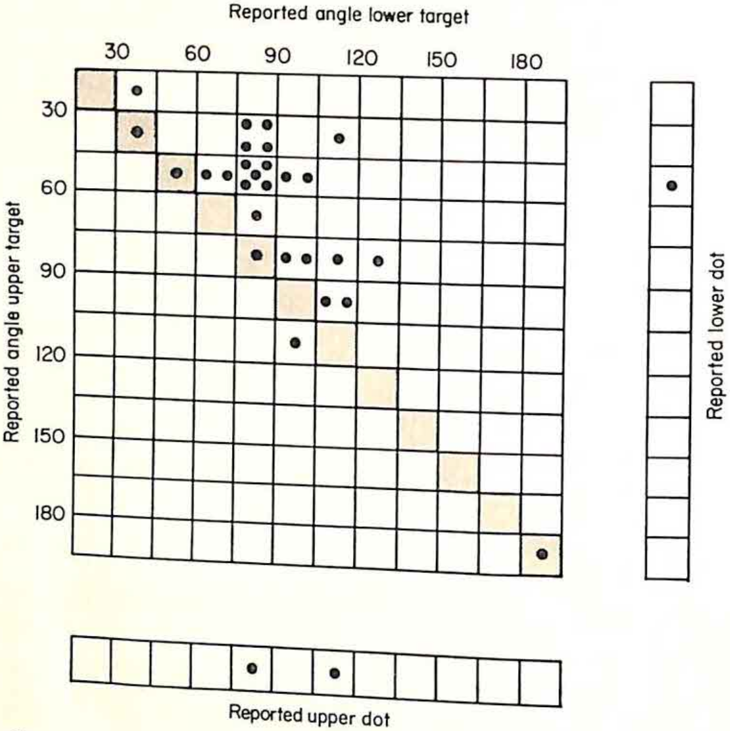


FIGURE 4. Results obtained from subject 1 in Experiment IV.

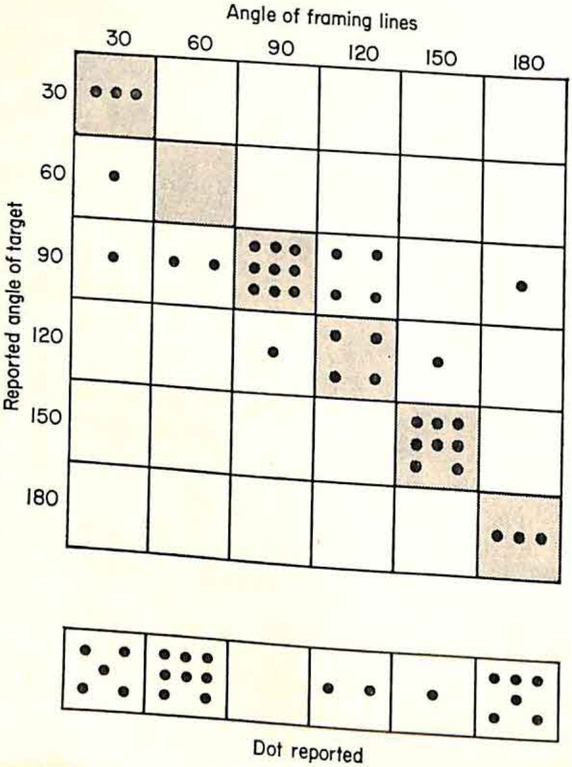


FIGURE 5. Results obtained from subject 3 in Experiment V.

targets had always appeared identical, all the points in Figure 4 would have fallen in the shaded squares. In fact, for only 4 out of 30 exposures was this the case. The apparent orientations of upper and lower targets could differ by as much as 82·5°; moreover, on 3 out of 30 exposures one target was perceived as a line, while the other was reported as a dot.

It is interesting that the preferred orientations for the upper and lower targets in this experiment differed by 30° despite the fact that Experiment III, in which the two targets were exposed separately, revealed no significant difference for the two sites of exposure. A similar small difference was obtained from one other subject. It seems possible that there is a genuine small difference in orientation preference for dot targets separated by 1·5°, but that the difference is so small that it is only made apparent in the conditions of Experiment IV, during which the two targets are shown simultaneously, one vertically above the other.

Similar results were given by the other four subjects and show that the orientations of two separate single-dot targets, exposed simultaneously, are frequently perceived as similar, but can be very different.

Experiment V: Can the preferred orientation for a given position in the visual field be changed?

We found that the preferred orientation for the thin lines observed when a dot is exposed at one point in the visual field, can be influenced by the presence of neighbouring straight lines. The subjects were tested with a single dot exposed briefly between two continuously visible and parallel black lines on the dimly illuminated white background provided by the pearl-glass screen. The orientation of the lines was arranged before each test, so that a variety of orientations was presented in random order. The lines were 3 mm. thick (8′ at the 128-cm. range of observation that we used), 8 cm. long and parallel; they were separated by 8 mm. (21′) and the dot was exposed midway between them. This central dot was exposed 10° from the optic axis, on the nasal side of the horizontal meridian.

Instructions to the subject were as follows: "In each trial you will be presented with a stimulus framed by two continuously visible lines. Please report the angle of the stimulus if it appears to be a line; tell us if you see a dot."

TABLE III

Subject	Lines perceived	Aligned reports	% Aligned reports
1	47	31	66·0
2	35	11	30·0
3	38	27	71·0
5	44	9	20·5
Expectation if no correlation			16·5%

The results for subject no. 3 are shown in Figure 5. The shaded squares in this figure show the places where all the points would fall, if the apparent orientation of the dot were always identical with that of the framing lines. Only 11 out of 38 of the targets perceived as lines appeared to have orientations that were different from the framing lines; in the case of this subject most of these differences occurred for reports of horizontal lines. It is also interesting that when the angle of the framing lines was 60°, the target was usually perceived as a dot. This observation is in good agreement with the polar diagram for the same subject (Figure 3) which shows that this subject perceived relatively few lines at 60°.

Table III summarizes the results for all four subjects who were given this test.

Discussion

The experiments described above were the outcome of an attempt to test a prediction. Perhaps 'prediction' is too strong a word; however, we have given our reasons in the introduction of this paper for believing that briefly exposed point-sources of light might appear as thin white lines to human observers if neurones in the human visual cortex show the same sort of orientation preference as do those of the cat and monkey. The results are consistent with this prediction and show that a point-source of light is indistinguishable from a thin white line in roughly two-third of tachistoscopic exposures. On the other hand, there are many aspects of the illusion described that cannot be explained.

The results of Experiments I and II show that small visual targets, consisting of either one, or a pair of point-sources of light, when exposed for near-threshold durations, are usually perceived as thin white lines. These lines may appear to be any length from 5′ to 30′ (2 to 12 mm. at the 128-cm. range of the screen) and may have any orientation. The two-dot target that we used was perceived as a thin white line in roughly two-third of exposures and the orientation of this line was often correctly reported as the actual orientation of the target (Figure 2); in the remaining trials with this target, a single dot or point-source of light was perceived (Table I). Exposure of a single dot during these trials was also perceived as a thin line in about two-third of the tests (Table I). This line could appear with any orientation and any length, but for one site in the visual field of a given subject, there was a preferred orientation (Figure 2). All of the subjects showed preferences near to the horizontal (Figure 3).

These are results that might be expected if neurones in the human visual cortex behave in a manner similar to those of the cat and monkey. It seems likely that near-threshold excitation of a small area of retina excites a relatively small number of neighbouring cortical cells. It would be the presumed orientation specificity of this small group of cortical neurones that might account for the illusion described by the results of our Experiments I and II. On the other hand, Andrews (1967) used optimal excitation of a small area of the retina in showing that each point in the field has a characteristic bias to both acuity for orientation and constant errors of perceived orientation. His short bright line stimulus may still have excited only a relatively small number of cortical cells. Hubel and Wiesel (1968) have shown that the orientation preference of cortical neurones in the monkey is not constant across the cortical surface. The visual cortex of this animal seems to be constructed as a functional mosaic of very small areas within which orientation preference is relatively constant; in one experiment the recorded orientation preference altered by 13° for a 20-μ shift of electrode position, measured parallel to the surface of the brain. Thus, if one assumes that similar neurophysiological functions operate in man, one might expect the orientation preference shown by our subjects exposed to dot-targets to vary from place to place in the visual field. The results of Experiment III are not consistent with this expectation. None of the five subjects showed any significant differences in orientation preference for dot-targets exposed at sites separated by 1° 15′ in the visual field.

Although the results of Experiments I and II were consistent with the conception of an illusion caused by strictly neurophysiological factors operating in the visual cortex, they might well be the consequence of quite different factors. For instance, they might be caused by astigmatic distortion of the image (Helmholtz, 1924) or by the small eye-movements of physiological nystagmus (Yarbus, 1967). Helmholtz discusses at some length the star-shaped image of a luminous point and the optical factors contributing to this aberration: these optical properties are clarified when later theories of light and optics are applied to image processing (Ditchburn, 1953). Experiments IV and V were designed to test the optical aberration explanation; with simultaneous presentation the numerous sources of aberration are maintained common for both stimuli.

The hypothesis that astigmatism is the cause of our observations is not supported by the results of either Experiments IV or V. One would expect that the astigmatic distortion of the images of two neighbouring and simultaneously exposed point-sources of light would be similar, and might depend only upon the state of accommodation of the eye at the time (Alpern, 1962). The results of Experiment IV do show that both targets are commonly perceived as lines with similar orientation, but these targets sometimes appeared to have quite different orientations. Moreover, there were times when one target appeared as a dot when the other was perceived as a line. Nor are the results of Experiment V consistent with the hypothesis of astigmatism, for it is hard to see how the presence of the two parallel framing lines used in this test could modify an astigmatic distortion. The hypothesis that the illusion we have described is the result of small eye-movements is not consistent with the results of Experiment IV; were movements of the eye during fixation the cause, one would have expected the appearance of two neighbouring targets to be invariably identical.

On the other hand, if one accepts orientation specificity of neurones in the human visual cortex as the most plausible explanation of our results, some explanation must be found for the following observations:

(i) when dot-targets are repeatedly exposed in the same part of the visual field, the orientations of the lines perceived are not always the same and can in fact differ by as much as 90°;

(ii) although there is a preferred or most likely orientation for these percepts, the preferred orientation does not vary when the target is presented in different parts of the visual field;

(iii) the preferred orientation can be controlled by the presence of straight lines near to the dot-target.

At present, we are inclined to believe that each individual percept is the consequence of the dot-target exciting a small number of cells in the visual cortex, which are orientation specific. Since these cells have always in the past been most readily excited by contrast borders with a particular orientation, their excitation in the conditions of these experiments is interpreted by the observer as indicating the 'customary' orientation. The different orientations reported in consecutive trials with the same target, would be due to inconstancy of image position during voluntary fixation causing the excitation of different groups of cortical cells on different occasions. Ditchburn (1956) has calculated that the image may wander within a circle of 25 to 50 cones diameter during voluntary fixation. This interpretation of our results forces one to conclude that the orientation specificity of human cortical units cannot be randomly distributed, in order to account for the preferred orientation of percepts reported by our subjects. Thus one would account for the 120° preference of subject no. 3 (Figure 2) by assuming *that the part of his visual cortex which we tested*, contained a preponderence of 120°-specific neurones. There is no independent evidence with which to support such a hypothesis: moreover, neither Hubel and Wiesel (1968) nor Campbell *et al.* (1968) could find any evidence for a non-random distribution of orientation specificities in the cortex of cat or monkey. Nevertheless, the idea appears consistent with the observation that human resolving power is not uniform for all orientations of a grating. Resolving power is significantly better for vertical and horizontal gratings than it is for oblique targets (Campbell, Kulikowski and Levinson, 1966). It is clear that some of the assumptions made above could be tested by presenting targets of the sort we have used as stabilized images.

The fact that the perceived orientation of lines sometimes deviates from the 'preferred' orientation may possibly be explained by assuming that the excitability of those cortical units affected by the presented dot pattern may change. These changes could either occur randomly due to spontaneous activity (Experiments II, III and IV), or systematically and predictably, due to the introduction of additional pattern elements into the visual field (Experiment V). From the results of his optimal viewing conditions study, Andrews (1967) has proposed that the relation between presented and perceived orientation is subject to adaptation and, also, that orientation-sensitive units in the human visual system integrate their outputs by a process of mutual inhibition.

Although the experiments described above have produced some surprising results, for some of which one can only offer tentative explanations, they seem to support the suggestion of Andrews (1967), Gilinsky (1967) and of Campbell and Kulikowski (1966), that neurones in the human visual cortex are orientation specific.

We wish to thank the Carnegie Institution of Washington for providing a Research Fellowship for G.M., and also the Canadian Defence Research Board for their support of R.P.

References

ALPERN, M. (1962). In DAVSON, H. (Ed.) *The Eye*, Vol. 3, pp. 191–229. New York: Academic Press.

ANDREWS, D. P. (1967). Perception of contour orientation in the central fovea. *Vision Res.* 7, 975–1013.

BARLOW, H. B., BLAKEMORE, C. and PETTIGREW, J. D. (1967). The neural mechanism of binocular depth discrimination. *J. Physiol., Lond.* **193**, 327–42.

BARLOW, H. B., HILL, R. M. and LEVICK, W. R. (1964). Retinal ganglion cells responding selectively to direction and speed of image motion in the rabbit. *J. Physiol. Lond.*, **173**, 377–407.

CAMPBELL, F. W., CLELAND, B. G., COOPER, G. F. and ENROTH-CUGELL, C. (1968). The angular selectivity of visual cortical cells to moving gratings. *J. Physiol., Lond.*, **198**, 237–50.

CAMPBELL, F. W. and KULIKOWSKI, J. J. (1966). Orientation selectivity of the human visual system. *J. Physiol., Lond.*, **187**, 437–46.

CAMPBELL, F. W., KULIKOWSKI, J. J. and LEVINSON, J. (1966). The effect of orientation on the visual resolution of gratings. *J. Physiol., Lond.*, **187**, 427–36.

DITCHBURN, R. W. (1953). *Light.* Glasgow: Blackie and Sons.

DITCHBURN, R. W. (1956). Eye movements in relation to retinal action. In *Problems in Contemporary Optics*, pp. 609–23. Arcetri-Firenze: Instituto Nazionale di ottica.

GILINSKY, A. S. (1967). Masking of contour detectors in the human visual system. *Psychonomic Sci.* **8**, 395–96.

HELMHOLTZ, H. (1924). In Southall, J. (Ed.) *Physiological Optics.* Menasha: Optical Society of America.

HUBEL, D. H. and WIESEL, T. N. (1959). Receptive fields of single neurones in the cat's striate cortex. *J. Physiol., Lond.*, **148**, 574–91.

HUBEL, D. H. and WIESEL, T. N. (1961). Integrative action in the cat's lateral geniculate body. *J. Physiol., Lond.*, **155**, 385–98.

HUBEL, D. H. and WIESEL, T. N. (1962). Receptive fields, binocular interaction and functional architecture in the cat's visual cortex. *J. Physiol., Lond.*, **160**, 106–54.

HUBEL, D. H. and WIESEL, T. N. (1963). Shape and arrangement of columns in cat's striate cortex. *J. Physiol., Lond.*, **165**, 559–68.

HUBEL, D. H. and WIESEL, T. N. (1968). Receptive fields and functional architecture of monkey striate cortex. *J. Physiol., Lond.*, **195**, 215–43.

YARBUS, A. L. (1967). *Eye Movements and Vision*, pp. 103–27. New York: Plenum Press.

(*At the request of the authors we have been asked to state that this paper was first received 3 March* 1969).

Q. Jl exp. Psychol. (1969) **21**, 312–321

MATCHING LATENCIES FOR WORD-SHAPE PAIRS

PHILIP H. K. SEYMOUR

Department of Psychology, The University, Dundee

Four word-shape pairs, *circle*/*CIRCLE*, *square*/*SQUARE*, *circle*/*SQUARE*, *square*/*CIRCLE*, were presented under conditions of simultaneous onset to 48 subjects. Subjects were instructed to report whether the pair could be classified as *same* or *different* by saying "yes" or "no" as rapidly as possible. Latencies of these judgements were measured from the simultaneous onset of the displays to the initiation of the subject's verbal response. The pair *circle*/*CIRCLE* was reported as *same* faster than the pair *square*/*SQUARE*. No consistent differences were obtained between the two mismatching pairs. However, latencies to report a mismatch were consistently longer than latencies to report a match.

Introduction

Oldfield (1966) refers to a study by Wingfield (1966) in which subjects were shown the printed name of an object followed by a picture of an object, and were instructed to say "yes" if the two were the *same* and "no" if they were *different*. Latencies, measured from the onset of the pictorial display to the initiation of the subject's response, were close to 500 msec. for objects having names of high and low frequency of usage. No difference was found between latencies to report a match as opposed to a mismatch.

Oldfield (1966) interpreted the absence of a name frequency effect as implying that the match/mismatch decision was not based on a nominal comparison. He also concluded that the results were consonant with the view that object naming was a two-phase process, consisting of a preliminary constant period, involving an estimate of familiarity and "any previous primary processes of reception, perception and so forth," and a search, of variable duration, for the object name. Oldfield was not explicit about the extent of categorization assumed to occur in the initial phase. The assumption of constant perception time across objects appears, however, to be at variance with the view adopted earlier by Oldfield and Wingfield (1965). In that paper it was proposed that object identification involved "progressive analysis of the presented characteristics of the object." This position, which is similar to that adopted by Bruner (1957), blurs the distinction between perceptual categorization and name retrieval, and admits the possibility of differences in classification time across object categories. Hence, if Wingfield's (1966) study is to be quoted as being in conflict with sequential models of the classification process it is important to assess the validity of his finding of an absence of variance in word-picture matching latencies.

The present paper reports two experiments which have been designed to explore more thoroughly certain limited aspects of the word-shape matching task.

These studies differ in some important methodological respects from those carried out by Wingfield. (1) The sequential presentation condition used by Wingfield has been replaced by a condition of simultaneous presentation of the pictorial and verbal displays. This change has been made because the factors of display duration and interdisplay interval are relevant parameters which require systematic examination (Davis, 1964; Corballis, Lieberman and Bindra, 1968). For the present, the assumption has been made that under conditions of simultaneous presentation all components of the overall latency will be compressed to minimum durations, and that the latency will therefore provide a sensitive indicator of changes in the operations of processing and comparison attributable to manipulation of display characteristics. The latencies reported here consequently include time to read the name as well as time to recognize the shape, and are expected to be somewhat longer than those reported by Wingfield. (2) A restricted ensemble of geometric shapes and their names has been selected in preference to the larger set of object names and naturalistic shapes used by Wingfield. This decision has been taken in order to reduce the problem of devising representations which are "modal" or "salient" (Brown, 1956). The ensemble consisted of four word-shape pairs generated by combining the names *square* and *circle* with the shapes *SQUARE* and *CIRCLE*. Within this ensemble factors of word length and familiarity are controlled. Both names specify rather precisely defined sets of two-dimensional shapes. Also, use of only two matching word-shape pairs facilitates control of stimulus and response probabilities.

The concepts *square* and *circle* can both be represented by a continuous perimeter line. However, the specifications required for the construction of *SQUARE* are more extensive than those required for the construction of *CIRCLE* (Attneave and Arnoult, 1956). The *CIRCLE* is an enclosed figure of high redundancy, whereas the *SQUARE* carries information critical for identification at each corner. In terms of the model considered by Oldfield and Wingfield (1965) it might then be argued that a smaller number of binary tests for the presence or absence of attributes are needed for classification of a shape as *CIRCLE* than for classification as *SQUARE*. Similarly, when the name *circle* is presented it may specify a shorter list of features criterial for inclusion in the class than the name *square*. On the assumption that word-shape matching involves a comparison of a list of features specified by a name and a list of features derived from analysis of a shape, it can be predicted that the pair *circle/CIRCLE* will be classified as *same* faster than the pair *square/SQUARE*. The prediction follows because the processes of (1) feature list generation, (2) shape analysis, and (3) comparison are all assumed to be less extensive for the concept *circle* than for the concept *square*.

The experiments also provide data on the times required to classify the mismatched pairs *circle/SQUARE* and *square/CIRCLE*. There is some evidence in the recent literature that more time is needed to report a mismatch than a match (Gough, 1965; Fillenbaum, 1964). Both mismatch pairs involve the comparison of a short feature list, derived from *circle* or *CIRCLE*, and a longer list, derived from *square* or *SQUARE*. In the absence of a more fully developed account of the comparison process a prediction of no difference between the mismatched pairs appears reasonable.

Experiment I

Method

Subjects

Twenty-four volunteers from classes in the Faculty of Social Science at the University of Dundee acted as subjects. All were naive as to the purposes of the experiment.

Apparatus

Two slide projectors projected down parallel tunnels onto a back-projection screen, divided to form two juxtaposed 8 × 8 in. display fields. When a switch was closed the mains power supply was passed through a resistance to the projector bulbs, bringing these to a red glow too weak to project a detectable image onto the screens. The same switch brought on a weak fixation light behind one or other of the two screens, which acted as a warning that display onset was imminent and as an instruction to fixate the centre of the illuminated screen. A second switch extinguished the warning/fixation light and brought both projector bulbs abruptly to full mains voltage, and at the same time started a Venner msec. stopclock. The timer was stopped by a signal from a voice-key mounted close to the subject, which also extinguished the projector bulbs. Slides were changed manually between each presentation.

Displays

Two sets of displays were prepared. The first consisted of the words *square* and *circle* printed in Letraset (Univers 55, 36 pt., lower case), and the shapes *SQUARE* and *CIRCLE* outlined in black ink. These displays were photographed and mounted as transparencies. In the second set word and shape were combined to form four displays in which the word was printed inside the shape, giving the combinations *square/SQUARE*, *circle/CIRCLE*, *square/CIRCLE* and *circle/SQUARE*. Subjects viewed the screens at a distance of 60 in. Under these conditions the shapes subtended a visual angle of 2 deg. 30 min., and the words a slightly smaller angle. The centres of the two screens were separated by 7 deg. of visual angle.

Procedure

Subjects were randomly allocated to one of three groups. All subjects saw 20 presentations of each of the four word-shape pairs, under the instruction to respond "yes" where the pair matched and "no" where the pair mismatched. Thus, frequencies of occurrence of the four stimulus pairs and of the two response categories were equated. The order of presentation was randomized and different for each subject, with the constraint that each word-shape pair should appear 5 times in each of 4 blocks of 20 presentations per block. Subjects were given initial practice in responding "yes" and "no" as a simple verbal reaction to illumination of the screens. They were then given practice in response to the word-shape pairs. The experiment was run as 4 blocks of 20 presentations, a short rest being given between each block. Subjects were instructed to respond as rapidly as possible consistent with the avoidance of error. Where errors did occur, or where the response failed to trigger the voice key, the presentation was subsequently reinserted in the random series.

Subjects in groups I and II responded to spatially separated displays, such that, for any given trial, the word was presented on one screen and the shape on the other. Subjects in these groups fixated the right-hand screen throughout two blocks of presentations, and the left-hand throughout the remaining two, two subjects in each group being allocated to one of the orders RLLR, LRRL, RLRL or LRLR. For subjects in group I the display appearing on the fixated screen might be either a shape or a word on a particular presentation, although the frequencies of the contingencies *fixate shape* and *fixate word* were equated within blocks and across word-shape pairs. For subjects in group II the class of display appearing on the fixated screen was consistent, half the subjects being allocated to a *fixate word* condition at blocks 1 and 2, and a *fixate shape* condition at blocks 3 and 4, this order being reversed for the remaining subjects. The distinction between groups I and II was, therefore, one between randomized and blocked scheduling of the *fixate word* and *fixate shape* conditions. Control of fixation was by instruction to fixate the centre of the screen illuminated by the warning light.

Subjects in group III saw the combined word-shape displays on a single screen throughout the experiment; manipulations with respect to direction of saccade and class of display first fixated were not applicable, therefore.

Subjects were tested individually in a quiet, darkened room. The test session lasted about 40 min.

Results

The error rate recorded was 3·8 per cent. Hence, 75 of the 1920 latencies recorded were repetitions following errors.

Mean response latencies for each subject under each of the four stimulus-response contingencies were calculated. The group means are summarized in Table I. These data were submitted to an analysis of variance within a three-factor design having repeated measures on two factors (Winer, 1962). The analysis indicated a significant effect for groups of subjects ($F = 22{\cdot}59$; *d.f.* = 2,21; $P < 0{\cdot}001$). Inspection of the data shown in Table I suggests that this effect is largely attributable to the faster response times of subjects in group III. Comparison of group means confirms that there is no difference between group I and group II ($t = 0{\cdot}3$; *d.f.* = 14; n.s.), but that group III differs significantly both from group I ($t = 6{\cdot}03$; *d.f.* = 14; $P < 0{\cdot}001$) and from group II ($t = 5{\cdot}79$; *d.f.* = 14; $P < 0{\cdot}001$). Thus, no difference was found for blocked vs. randomized scheduling of the *fixate word* and *fixate shape* contingencies. The combined presentation condition was, however, associated with significantly faster response times, the difference being of the order of 130 msec.

TABLE I

Response latencies (msec.) to matched and mismatched word-shape pairs (Experiment I)

Response Display	"Yes" circle/ CIRCLE	 square/ SQUARE	Mean match	"No" circle/ SQUARE	 square/ CIRCLE	Mean mismatch
Group I spatial separation	748	768	758	792	823	808
Group II spatial separation	720	765	743	816	801	809
Group III displays super-imposed	593	634	614	688	665	677

Means for each display pair are based on 160 readings.

Factors examined within subjects were response category (match/mismatch) and word presented (*square/circle*). A highly significant effect was obtained for response category ($F = 90{\cdot}29$; *d.f.* = 1,21; $P < 0{\cdot}001$). It is clear from Table I that this effect arises because latencies to report a match are consistently faster than latencies to report a mismatch. The difference is of the order of 60 msec.

The main test of the simple prediction outlined in the Introduction was made by examining the effect of the word presented. This effect is also highly significant ($F = 9.18$; $d.f. = 1,21$; $P < 0.01$). However, a significant interaction was obtained between word displayed and response category ($F = 8.63$; $d.f. = 1,21$; $P < 0.01$). Inspection of the data supports the view that the effect for word presented arises because the pair *circle*/*CIRCLE* is associated with faster latencies than the pair *square*/*SQUARE*; the interaction occurs because comparable differences were not obtained between the two mismatch pairs.

The validity of this interpretation was tested by carrying out comparisons among the ordered means of the display pairs for each group, following the Newman–Keuls procedure (Winer, 1962). The comparisons showed that the pair *circle*/*CIRCLE* was associated with faster response times than the pair *square*/*SQUARE* for all groups, and that this difference was significant at, at least, the 0·05 level for groups II and III but not for group I. No significant differences were obtained between the mismatched pairs, *square*/*CIRCLE* and *circle*/*SQUARE*. However, both matching pairs differed significantly from both mismatched pairs, with the exception of group I, where the difference between *square*/*SQUARE* and *circle*/*SQUARE* was not significant.

The results of Experiment I were, in general, in line with prediction, therefore, in showing (1) faster times for the pair *circle*/*CIRCLE* than for the pair *square*/*SQUARE*, (2) faster times for a match than for a mismatch, and (3) no difference between the mismatch pairs *square*/*CIRCLE* and *circle*/*SQUARE*. In the absence of interactions with groups of subjects it can be concluded that the effects of concept and decision outcome are not influenced by the spatial characteristics of the displays or blocked or randomized scheduling of the *fixate word* and *fixate shape* contingencies. However, the comparisons among means showed some attenuation of these effects for group I (randomized scheduling). Inspection supported the view that, for group I, the *fixate shape* contingency was slower than the *fixate word* contingency, although this effect did not occur for group II (blocked scheduling). One interpretation of this finding might be that the *fixate word* contingency corresponds to the preferred strategy in word-shape matching, and that the blocked schedule facilitates adoption of this strategy, whereas randomized scheduling tends to disrupt it. The condition applied to group I appears, therefore, to raise issues concerning the order of input of information in the word-shape matching task which are outside the scope of the present paper. For this reason, the decision was taken to run a second experiment designed to confirm the validity of the findings for the blocked scheduling condition in Experiment I.

Experiment II

Method

Subjects

These were 24 volunteers from a class in education and psychology at the University of Dundee, all of whom were naive as to the aims of the experiment.

Apparatus and displays

These were as in Experiment I, with the exception that the combined word-shape displays were not used.

Procedure

Subjects were initially allocated to one of four groups. The experiment differed from Experiment I in that the factors of direction of saccade and class of display fixated were placed between groups. For block 1, consisting of 48 presentations, subjects in groups I and II fixated the right-hand screen, and subjects in groups III and IV fixated the left-hand screen. Groups I and III were allocated to a *fixate shape* condition, and groups II and IV to a *fixate word* condition. Control of the direction of saccade was again by instruction to fixate the centre of the screen showing the warning light immediately before display onset. Each of the four word-shape pairs appeared 12 times in the block in an order which was random and different for each subject.

Following completion of block 1 a second block of 48 presentations was given. Half the subjects in each group switched at this point from the *fixate word* to the *fixate shape* condition, or vice versa. The remaining subjects were allocated to the same condition under block 2 as under block 1. The design therefore permitted a comparison, across two blocks of 48 presentations, of *fixate shape/fixate shape*, *fixate word/fixate word*, *fixate shape/fixate word* and *fixate word/fixate shape* conditions, and represented an extension of the procedures applied to group II in Experiment I.

Results

The error rate recorded was 3·99 per cent, or 92 repetitions following error out of a total of 2304 latencies.

A preliminary analysis was carried out on the mean response latencies for each subject under each word-shape pair at block 1 within a three-factor design with repeated measures on one factor (Winer, 1962). No effects for the between group factors of direction of saccade or class of display fixated were obtained, F ratios being less than 1 in each case. A significant effect was obtained for display pairs ($F = 35{\cdot}74$; $d.f. = 3,60$; $P < 0{\cdot}001$). No interactions were significant.

In the absence of an effect for direction of saccade this factor was dropped from the analysis, and subjects were regrouped on the basis of the class of display (word or shape) fixated under blocks 1 and 2, giving four groups, corresponding to the conditions *shape/shape*, *word/word*, *shape/word* and *word/shape*. The data for these four groups are summarized graphically in Figure 1. These data were submitted to analysis of variance within a three-factor design, with repeated measures on two factors (blocks and display pairs). The analysis confirmed the absence of significant differences between groups of subjects ($F < 1$; $d.f. = 3,20$; n.s.). The effect of blocks of presentations was also not significant ($F = 3{\cdot}61$; $d.f. = 1,20$; $P < 0{\cdot}1$), and no interaction between groups and blocks was obtained ($F = 1{\cdot}44$; $d.f.$ 3,20; n.s.). This provides support for the conclusion based on examination of the data for group II in Experiment I: under conditions of blocked scheduling no variance in response times is attributable to use of the *fixate word* and *fixate shape* contingencies, or to a shift from one contingency to the other.

The effect of display pairs was again highly significant ($F = 50{\cdot}16$; $d.f. = 3,60$; $P < 0{\cdot}001$). Inspection of Figure 1 suggests that this effect is attributable to slower response latencies under the mismatch contingency, and to faster response times to *circle/CIRCLE* than to *square/SQUARE* under the match contingency. The Newman–Keuls procedure was used to carry out comparisons among the display means for each group at each block of presentations. Mismatched pairs were associated with significantly slower response times than matching pairs for

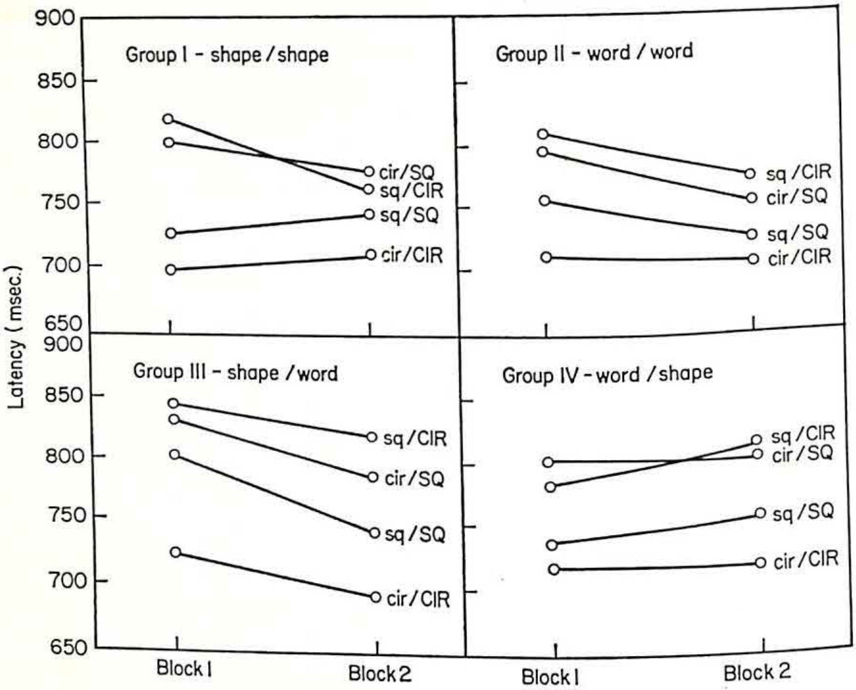


FIGURE 1. Response latencies (msec.) to matched and mismatched word-shape pairs (Experimen II).

N.B. Each point is the mean of 72 observations. Open circles indicate the condition *fixate shape* and filled circles the condition *fixate word*.

all groups at both blocks. The pair *circle/CIRCLE* was faster, at, at least the 0·05 level, than the pair *square/SQUARE* for all comparisons made. Trends were again less obvious under the mismatch contingency. However, the pair *circle/SQUARE* was significantly faster than *square/CIRCLE* for group I at block 1, and groups II and IV at block 2, the order being reversed for group IV at block 1.

Experiment II consequently serves to confirm the findings of Experiment I.

General Discussion

The analyses of the two experiments which have been reported provide general support for the simple hypothesis which was outlined in the Introduction, and raise a number of additional issues. The initial hypothesis proposed was that, in a situation where a subject must test a presented pattern for inclusion in a class of patterns specified by a verbal label, the speed with which the outcome of this test can be reported will be related to the length of the list of features which are critical for placement of a pattern in the class. It was argued *a priori* that the label *circl* specifies a shorter list of features than the label *square*. The experimental findin that the pair *circle/CIRCLE* can be accepted as congruent more rapidly than th pair *square/SQUARE* is compatible with this position. Since the two verba labels are matched for familiarity, letter length and frequency of occurrence in the experiment, and since both pairs are associated by instruction with the same response, it seems that the difference must be located among the inferred processes of feature list specification, shape analysis, or comparison.

These conclusions are based on the assumption that the word-shape matching task is logically subdivisible into four subordinate operations, corresponding to the processing of the two displays, a decision or comparison process, and a process of response selection and initiation. Figure 2 represents this simple breakdown in diagramatic form. Following the type of hierarchical task analysis proposed by Gagne (1962) and Miller, Galanter and Pribram (1960) the main operations have been elaborated to indicate their dependence on the completion of a number of subroutines, each of which could be further analysed. This analysis is presented merely as an aid to discussion of the experimental task. Further, although the operations identified appear logically distinct, they need not be treated as temporally distinct.

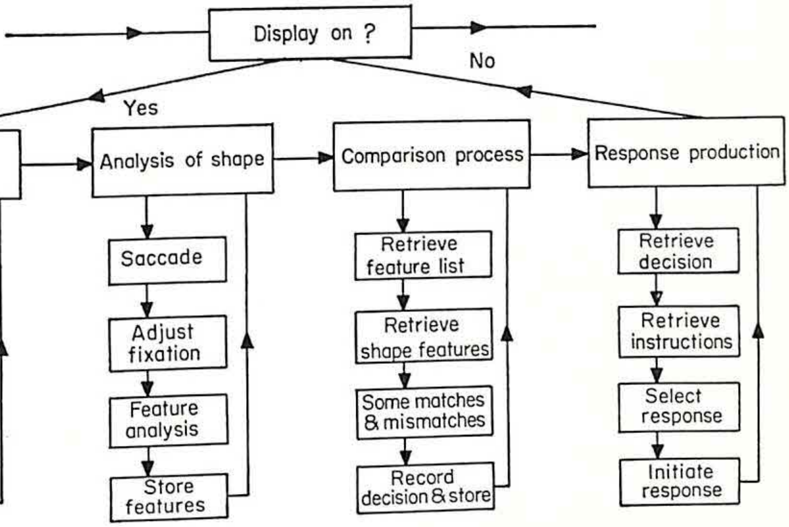


FIGURE 2. Task analysis of operations involved in word-shape matching.

The data obtained from the present experiments may be interpreted as differences in the durations of the processes of feature derivation, analysis and comparison between the word-shape pairs *circle/CIRCLE* and *square/SQUARE*. If this interpretation is valid it follows that variance in word-shape matching times may be expected which will be related, on the input side, to such factors as the specificity of the label, the characteristics of the class of patterns designated, and the complexity of the presented shape, but not necessarily to the general frequency of usage of the label. This view implies a close interdependence between nominal and classificatory processes, and raises some question as to whether Wingfield's (1966) prototype matching task should be taken as evidence for the functional independence of recognition and naming. However, much depends on the degree to which the present findings for geometric shapes can be generalized to naturalistic shapes, and on the distinction between sequential and simultaneous presentation conditions.

The second major finding was that latencies were consistently longer where the word and shape did not match than where they did. Since a difference in response category is involved here, it is possible that this finding is a simple artifact of the characteristics of the vocalizations "yes" and "no". This view is not supported by an examination of times to say "yes" or "no" as simple verbal reactions to illumination of the display screens. Mean times obtained from 24 unpractised subjects were 314 msec. for "yes" and 312 msec. for "no". These times represent the duration of the operations listed under Response production in Figure 2. There appears to be no reason to locate the match/mismatch difference within the processes of shape and word analysis, since these are logically prior to the process of comparison, on which the selection of the response depends. The inference is, then, that the match/mismatch difference relates to some characteristic of the comparison process, and that the processes leading to the output *mismatch* differ from those leading to the output *match*. Bindra, Williams and Wise (1965), discussing the opposite result in a study of matching tones of low discriminability, interpret their findings in terms of two statistical adders, one accumulating information about similarities, the other about dissimilarities. A view of this type would be consonant with the notion of feature list matching which has been adopted here, given the additional assumption that, under conditions of high interdisplay discriminability, the criterion for the decision *same* will be set lower than the criterion for the decision *different*.

Brief reference may also be made to the findings for eye-movements and the spatial characteristics of the display. Latencies were considerably shorter where both word and shape were presented simultaneously within foveal vision than where they were separated by 7° of visual angle. Gilbert (1959) has also reported decrements in performance on word recognition tasks involving saccadic movements of the eyes. The decrement may relate to processes of adjustment during refixation Tinker (1958), and the saving of 120 to 130 msec. obtained here may be simply an estimate of the duration of the saccade and refixation. On the other hand, subjects operating under the combined display condition in Experiment I were not forced to take two looks, one at the word and one at the shape, and it is possible that they were reading all the necessary information off the display within a single fixation and were then proceeding to process the word and shape in parallel. No effect was obtained for direction of saccade in Experiment II. This factor was not analysed in Experiment I, but inspection confirmed the absence of a difference. It remains possible, however, that directional reading habits may be important in a situation of this type (Heron, 1957; Bryden, 1967), and further experimental study of this factor is required.

Inspection of the data for Group I in Experiment I suggested that the *fixate shape* condition might be associated with slower times than the *fixate word* condition. If this finding is supported in subsequent studies it would provide support for subjects' introspective reports that first fixation of the word represents the most efficient strategy in this situation. Subjects operating under conditions of consistent scheduling may simply adapt their eye-movements to achieve first fixation of the word irrespective of the class of display appearing on the fixated screen.

References

ATTNEAVE, F. and ARNOULT, M. D. (1956). The quantitative study of shape and pattern perception. *Psychol. Bull.* **53**, 452–71.

BINDRA, D., WILLIAMS, J. A. and WISE, J. S. (1965). Judgements of sameness and difference: Experiments on decision time. *Science, N.Y.* **150**, 1625–7.

BROWN, R. (1956). Language and categories. In BRUNER, J. S., GOODNOW, J. J. and AUSTIN, G. A. (eds), *A Study of Thinking*. New York: John Wiley.

BRUNER, J. S. (1957). On perceptual readiness. *Psychol. Rev.* **64**, 123–52.

BRYDEN, M. P. (1967). A model for the sequential organisation of behaviour. *Canad. J. Psychol.* **21**, 37–56.

CORBALLIS, M. C., LIEBERMAN, W. and BINDRA, D. (1968). Discriminability and central intermittency in same-different judgements. *Q. Jl exp. Psychol.* **20**, 51–61.

DAVIS, R. (1964). The combination of information from different sources. *Q. Jl. exp. Psychol.* **16**, 332–9.

FILLENBAUM, S. (1964). Semantic satiation and decision latency. *J. exp. Psychol.* **68**, 240–44.

GAGNÉ, R. M. (1962). The acquisition of knowledge. *Psychol. Rev.* **69**, 355–65.

GILBERT, L. C. (1959). Saccadic movements as a factor in visual perception in reading. *J. educ. Psychol.* **50**, 15–19.

GOUGH, P. B. (1965). Grammatical transformations and speed of understanding. *J. verb. Learn. verb. Behav.* **4**, 107–11.

HERON, W. (1957). Perception as a function of retinal locus and attention. *Am. J. Psychol.* **70**, 38–48.

MILLER, G. A., GALANTER, E. and PRIBRAM, K. H. (1960). *Plans and the Structure of Behaviour*. New York: Holt, Rinehart and Winston, Inc.

OLDFIELD, R. C. (1966). Things, words and the brain. *Q. Jl exp. Psychol.* **18**, 340–53.

OLDFIELD, R. C. and WINGFIELD, A. (1965). Response latencies in naming objects. *Q. Jl exp. Psychol.* **17**, 273–81.

TINKER, M. A. (1958). Recent studies of eye movements in reading. *Psychol. Bull.* **55**, 215–31.

WINER, B. J. (1962). *Statistical Principles in Experimental Design*. New York: McGraw-Hill.

WINGFIELD, A. (1966). *The Identification and Naming of Objects*. D. Phil. Thesis, Oxford.

Received 10 *June*, 1969

Q. Jl exp. Psychol. (1969) **21**, 322–329

THE ROLE OF THE CS AND AVOIDANCE PERIOD DURATION IN DISCRIMINATED AVOIDANCE CONDITIONING

GERALD B. BIEDERMAN
University of Toronto

In an experiment investigating the effect of CS duration on discriminated bar-press conditioning, subjects were assigned to one warning period duration (1·5, 5·0 or 15·0 sec.) on the first day of avoidance training, and to one of the three durations on the second day of training. On each day avoidance behaviour was greatly influenced by the duration of the CS (warning) period, but the duration on day I had no effect on the second day's avoidance performance. Groups receiving no CS during training, although provided an avoidance contingency, showed little conditioning, but produced highly significant amounts of intertrial responding. In a second experiment, subjects receiving CS on day I were shifted to no CS on day II. Avoidance performance on day II was not significantly different from the day II performance of subjects in Experiment I having two training sessions with the CS present or absent on both days.

Studies of the stimulus control of discriminated avoidance conditioning have been largely concerned with onset and offset of the CS relative to the appearance of the US, following the Mowrer and Lamoreaux (1942) experiment which employed groups for whom CS terminated before, with, or after the CR (e.g. Kamin, 1956; Ulrich, Holz and Azrin, 1964; and most recently, D'Amato, Fazzaro and Etkin, 1968).

A parameter which has not enjoyed extensive experimental investigation in the Skinner box situation is the duration of the warning period itself. Bolles (1967, p. 411) concluded, mainly from shuttle-box evidence, that warning period duration had its effect in the maintenance rather than the acquisition of discriminated avoidance. Pearl and Edwards (1962) performed a Skinner box study which indicated both an acquisition and a maintenance effect, depending upon the method employed to shift subjects from one CS duration to another.

The present experiments assess the role of the duration of the warning period in avoidance acquisition, and attempt to determine the function of the CS itself by training avoidance behaviour with an avoidance period in force, but with no CS provided.

Experiment I

The design of this experiment is similar to the factorial design cited by Kimble (1961, p. 119) which had the purpose of separating learning and performance effects.

Method

Subjects

Fifty-four male Sprague-Dawley rats approximately 100 days old at the beginning of the experiment, served as subjects. The subjects were experimentally naive.

Apparatus

The apparatus consisted of two commercial (Grason-Stadler) two-bar experimental stations with the right bar removed. Scrambled foot shock from a Grason-Stadler shock generator (E6070B) was delivered to the subject through the floor grid with a shock intensity of 1·0 mA. The CS was white noise, while the ventilating fan provided a masking noise within the enclosures containing the experimental stations. The fan noise measured 70 dB at the rear of the chamber, with the CS bringing the sound intensity to 80 dB. All programming was automatic; responses were recorded by an Esterline-Angus event recorder and by electro-mechanical counters.

Procedure

On the day preceding training the escape response was shaped (as necessary); this response consisted of a bar-press during shock. The avoidance contingency was in force, but avoidance responding during the shaping trials was rare. The shock used in shaping (trials 1 to 20) was continuous shock terminated by the escape response; on the remaining trials (21 to 40), following a failure to avoid, the shock cycled 0·2 sec. on, 2·0 sec. off, until the escape response was produced (D'Amato, Keller and DiCara, 1964).

Twenty-seven subjects were randomly selected to receive CS (white noise), while the remaining 27 subjects received no warning stimulus, although the avoidance contingency was available to them. The CS group and the group with no CS were randomly assigned, respectively, to one of three avoidance interval durations (1·5, 5·0 or 15·0 sec.) for the first day of training and to one of the three durations for the second day of training, in a 3 × 3 factorial design. The avoidance interval for a subject on day I of training was used on the shaping session. Subjects received 500 trials each day; each trial consisted of the avoidance period, in which a response prevented the occurrence of shock and terminated the trial, or in the event of a failure to avoid, the avoidance period was followed by discontinuous shock which terminated with the escape response. For CS groups the white noise coincided with the beginning of each trial and was terminated when the trial ended. Each response (avoidance or escape) was followed by a VI 45-sec. intertrial interval (ITI) programmed by a film-tape programmer. Bar presses during the ITI were recorded but they had no further consequences.

Results

Avoidance conditioning was measured by the mean percentage of avoidance on the last 400 trials of each day of training (D'Amato, Keller and Biederman, 1965). In addition the general operant level was measured by the mean intertrial responses (ITR) per 100 trial blocks, for the final 400 trials of training (Biederman, 1967). The main result of Experiment I is given in Figure 1 for CS groups and in Figure 2 for no CS groups. The day I warning period duration had no significant effect on day II performance ($F < 1$) while the day I performance difference (CS group Figure 3), as a function of warning period, were significant ($F = 7{\cdot}09$, $d.f. = 2, 48$; $P < 0{\cdot}005$); the day II performance differences as a result of warning duration on day II were similarly significant ($F = 19{\cdot}67$, $d.f. = 2, 36$; $P < 0{\cdot}005$).

The presence of the CS increased avoidance performance on both days I and II; F (day I) $= 16{\cdot}60$, $d.f. = 1, 48$; $P < 0{\cdot}005$; F (day II) $= 31{\cdot}25$, $d.f. = 1, 36$;

$P < 0{\cdot}005$; (Figures 1, 2 and 3). No interactions from the analysis of days I and II data, respectively, were significant ($F < 1$, in all cases).

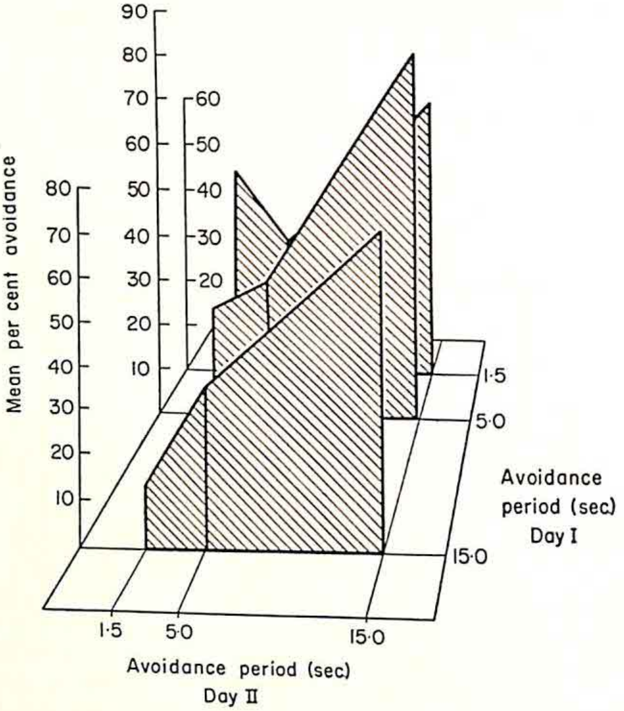


FIGURE 1. Avoidance performance on last 400 trials on day II of training with CS present.

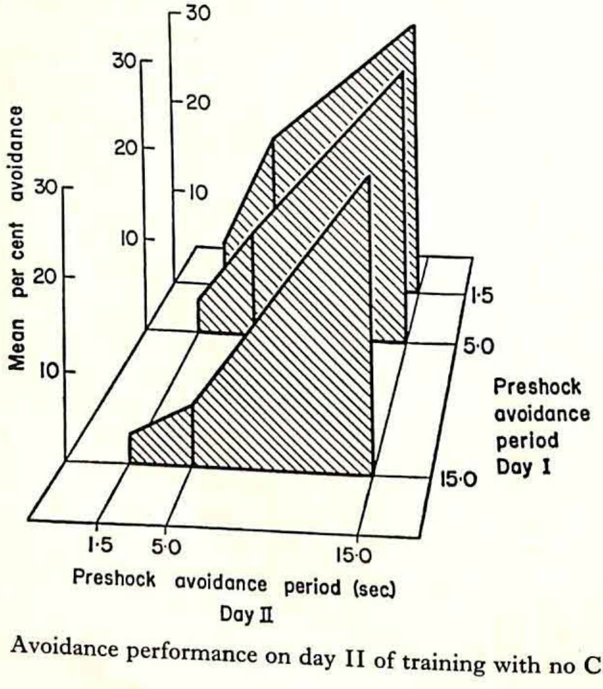


FIGURE 2. Avoidance performance on day II of training with no CS present.

Analysis of intertrial responding for both days I and II reveals that the absence of the CS greatly affected the amount of responding during the ITI. The no CS group produced approximately three times the level of ITR when compared with the CS group's rate (Figure 4); the no CS group's rate of response doubled by day II, as shown in Figures 5 and 6 F (day I) $= 35{\cdot}43$, $d.f. = 1, 48$; $P < 0{\cdot}005$; F (day II) $= 99{\cdot}86$, $d.f. = 1, 36$; $P < 0{\cdot}005$). No other main effect or interaction on days I or II reached significance (in each case, $F < 1$).

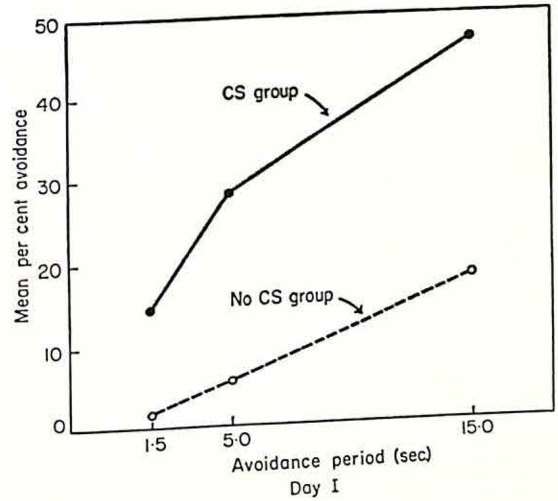


FIGURE 3. Mean per cent avoidance on day I of training.

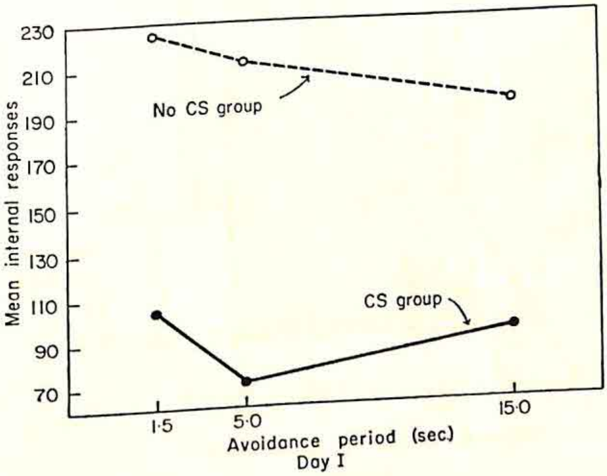


FIGURE 4. Mean intertrial responses on day I of training.

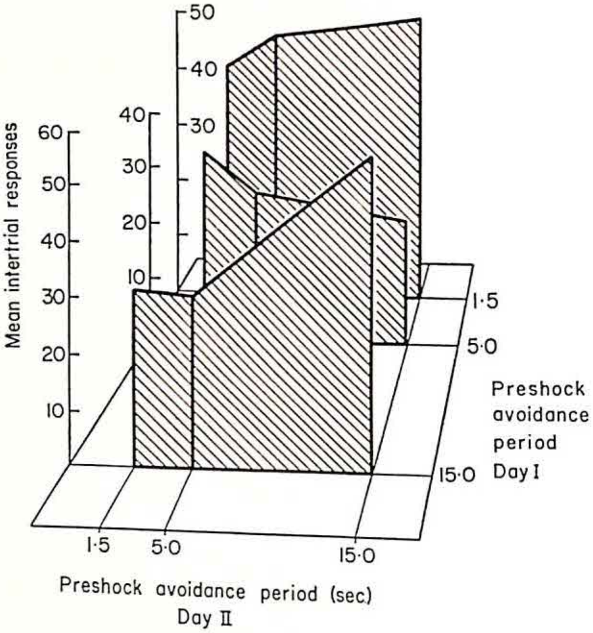


FIGURE 5. Mean intertrial responses per 100 trials over last 400 trials on day II of training, with CS present.

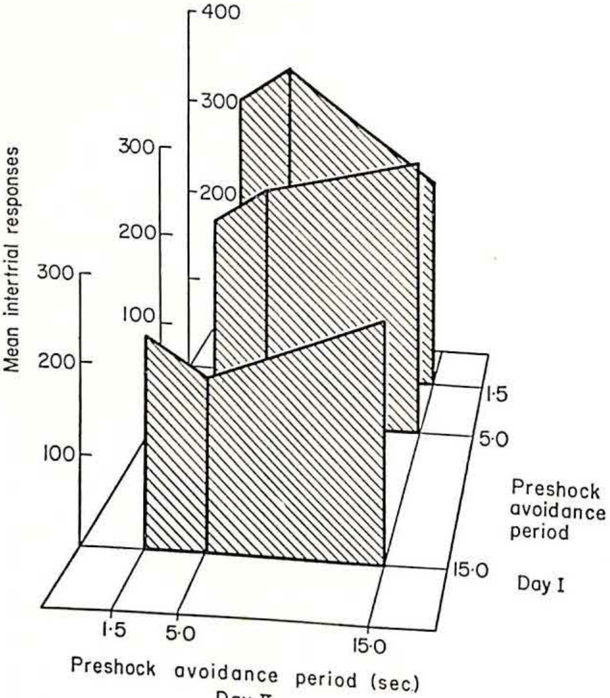


FIGURE 6. Mean intertrial responses day II with no CS present.

Experiment II

The purpose of this experiment was to show the contribution of the presence of the CS to the acquisition of avoidance behaviour, by changing the CS condition within groups.

Method

Subjects

Twenty Sprague-Dawley males, 100 days old, served as subjects.

Apparatus and procedure

The apparatus and general procedure were similar to Experiment I. In the present experiment all subjects received a 15·0 sec. avoidance period. Ten subjects were randomly selected to receive CS (white noise) on the first day of training and no CS on the second day. Ten subjects received no CS on the first day of training and received CS on the second day; within each group day I and shaping stimulus conditions were identical. All other experimental parameters were the same as in Experiment I.

Results

A summary of the results of Experiment II is given in Table I. The group initially receiving CS on day I and then shifting to no-CS on day II, shows no significant decrement in performance (correlated $t < 1$), whereas the group originally without CS and then shifted to CS shows a significant improvement (correlated $t = 2{\cdot}67$, $d.f. = 9$, $P < 0{\cdot}05$). When the day II performance in Experiment II for the no CS—CS group is compared with the group in Experiment I that had a 15-sec. warning period and CS on both days no significant difference is revealed ($t = 1{\cdot}00$, $d.f. = 17$). Similarly, when the CS—No CS group is compared to the group in the first experiment which had 15-sec. avoidance training for both training days but no CS on either day, the difference in the percentage avoidance is not significantly different ($t < 1$, $d.f. = 17$).

TABLE I

Mean percentage of avoidance responses (Experiment II)

Group	Day I (%)	Day II (%)
CS, No CS	37·38	31·60
No CS, CS	30·38	59·50

Discussion

The results of the first experiment support Bolles' (1967) conclusion that warning period duration affects performance rather than avoidance learning. The fact that subjects who are given greater opportunity to learn (and who actually perform at a higher avoidance level) fail to show subsequent superiority when their second-day performance is considered, requires explanation.

It is possible that during longer avoidance periods the initial fear response to the onset of the CS may be given more chance to dissipate and not interfere with the production of the avoidance response, but why this should not result in savings for groups remaining at the 15-sec. duration for the two training days is not clear.

The intertrial responding in Experiment I presents at first glance a similarly puzzling state of affairs. Subjects without CS but who none-the-less have the opportunity to avoid shock produce far more intertrial responses than subjects who receive CS throughout conditioning. If intertrial responding reflects the general operant level the probability then exists that subjects with a higher operant level should be more likely to respond during the avoidance interval. It is more reasonable to assume, however, that the intertrial responses are a consequence of the occurrence of the shock itself. Without the temporal cues provided by the CS, and with a variable intertrial interval, the occurrence of the shock may produce a considerable startle-response which may in turn produce a burst of bar-pressing. To this extent the CS may serve as a fear reducing factor (Rescorla and LoLordo, 1965).

The results of Experiment II may be interpreted to suggest that the CS has a more important role in the maintenance rather than the acquisition of avoidance behaviour. If avoidance learning in the absence of the CS were not taking place the day II performance of the group shifted from no CS to CS should be at about the level of the day I performance of the group receiving CS on day I, however, the former group showed considerably more avoidance behaviour on day II. The acquisition of the avoidance operant may arise simply through the accidental response in the "warning" period and subsequent reinforcement through the avoidance of shock. The emitting of the operant may be controlled by the temporal cues provided by the CS which might be better characterized as a performance rather than an acquisition role (Sidman, 1966).

In summary, it may well be the case that the effects of varying the duration of the warning period pose a more complex theoretical problem than the control of avoidance responding by the CS which has been apportioned the lion's share of experimentation and explanatory consideration.

This research was supported by grants from the National Research Council of Canada (no. 185) and from the Ontario Mental Health Foundation (no. 106). The experiment was performed by G. Heighington and B. Heighington.

References

Biederman, G. B. (1967). Discriminated avoidance conditioning: Stimulus function in shaping and training. *Psychon. Sci.* **9**, 263–4.

Bolles, R. C. (1967). *Theory of Motivation.* New York: Harper & Row.

D'Amato, M. R., Fazzaro, J. and Etkin, M. (1968). Anticipatory responding and avoidance discrimination as factors in avoidance conditioning. *J. exp. Psychol.* **77**, 41–7.

D'Amato, M. R., Keller, D. and Biederman, G. (1965). Discriminated avoidance learning as a function of parameters of discontinuous shock. *J. exp. Psychol.* **70**, 543–8.

D'Amato, M. R., Keller, D. and DiCara, L. (1964). Facilitation of discriminated avoidance learning by discontinuous shock. *J. comp. physiol. Psychol.* **58**, 344–9.

KAMIN, L. J. (1956). The effects of termination of the CS and avoidance of the US on avoidance learning. *J. comp. physiol. psychol.* **49**, 420–4.

KIMBLE, G. A. (1961). *Hilgard and Marquis' Conditioning and Learning*. New York: Appleton-Century-Crofts.

MOWRER, O. H. and LAMOREAUX, R. R. (1942). Avoidance conditioning and signal duration—a study of secondary motivation and reward. *Psychol. Monogr.* **54** (5, Whole No. 247).

PEARL, J. and EDWARDS, R. E. (1962). Delayed avoidance conditioning: warning stimulus (CS) duration. *Psychol. Rep.* **11**, 375–80.

RESCORLA, R. A. and LOLORDO, V. M. (1965). Inhibition of avoidance behaviour. *J. comp. physiol. Psychol.* **59**, 406–12.

SIDMAN, M. (1966). Avoidance behaviour. In Honig, W.K. (ed.), *Operant Behaviour; Areas of Research and Application*, pp. 448–498. New York: Appleton-Century-Crofts.

ULRICH, R. E., HOLZ, W. C. and AZRIN, N. H. (1964) Stimulus control of avoidance behaviour. *J. exp. Anal. Behav.* **7**, 129–33.

Received 10 *June* 1969

Q. Jl exp. Psychol. (1969) **21**, 330–338

DISCRIMINATION OF CIRCLES AND SQUARES WITH AND WITHOUT KNOBS BY GOLDFISH

N. S. SUTHERLAND AND ROSEMARY BOWMAN

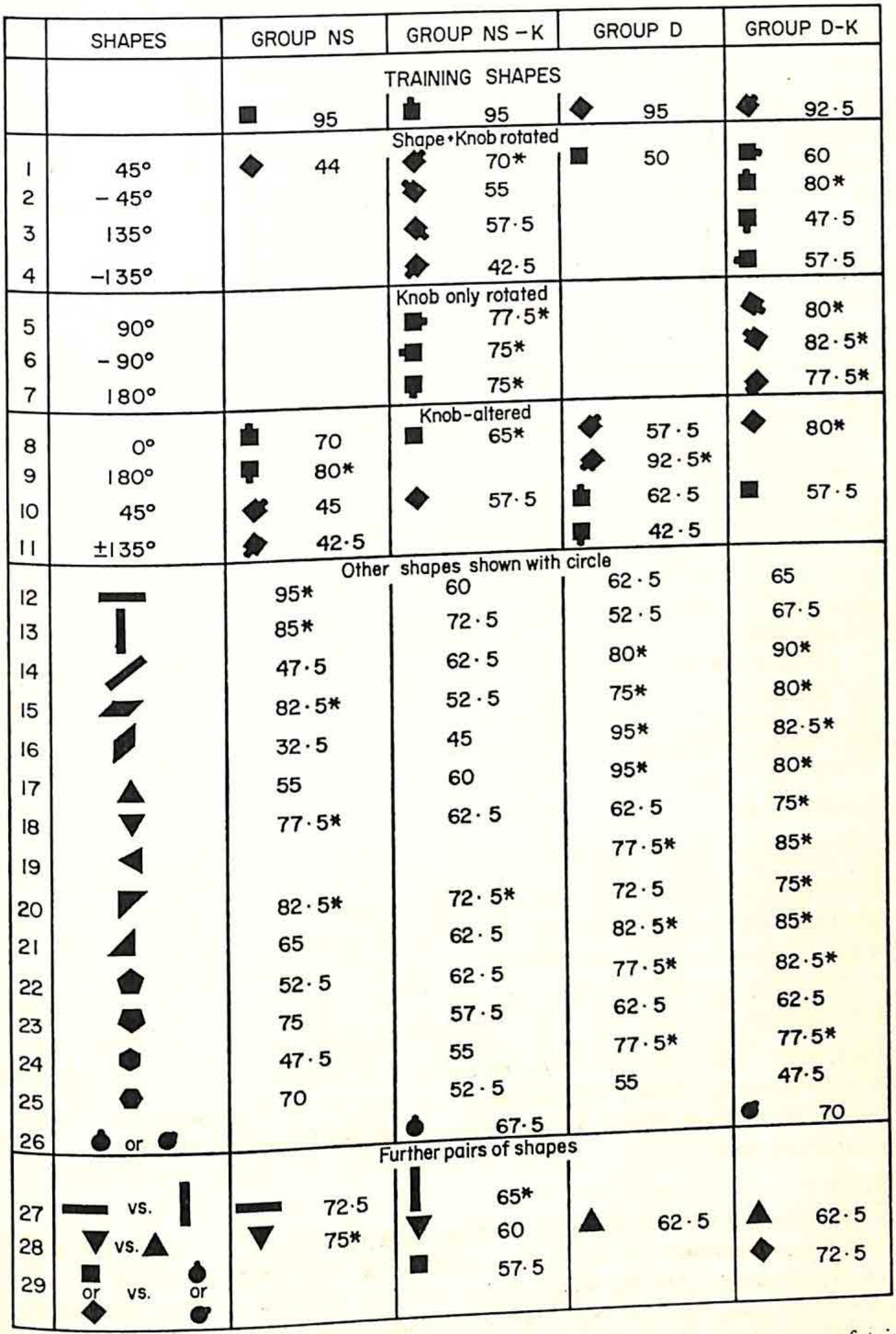
Laboratory of Experimental Psychology, University of Sussex

Different groups of goldfish were trained to discriminate between a circle and either a square, a diamond, a square with a knob added at the top or a diamond with a knob. All subjects readily mastered the discrimination, and there was no significant difference between the groups in the number of trials taken to criterion. In transfer tests it was found that: (i) while there was no transfer from square to diamond or vice versa, the presence of knobs on these shapes did mediate transfer from one to the other, provided the knob remained in the top part of the shapes; (ii) all groups learned much more about the tops of the shapes than about the bottom halves; (iii) animals discriminated the square from the circle largely by detecting the presence of a strong horizontal edge at the top of the figure; (iv) two features of the diamond were discriminated—the presence of oblique contours and the existence of a point in the upper half of the shape. The results are discussed in terms of a proposed model of shape recognition.

Introduction

Several theoretical accounts of shape discrimination have assumed that shapes remain invariant under rotation (e.g. Deutsch, 1955). Existing evidence (see for example Sutherland, 1963) shows that after animals have been trained to discriminate between two shapes, some rotated shapes are classified as the same as the original shape, others are not. We know remarkably little about what determines whether a shape will remain invariant for an animal under rotation. Sutherland (1968) has proposed a theory of shape recognition in animals which predicts that rotational invariance will only be found in special cases. According to this theory, the description of a shape stored in the animal's brain is highly abstract. Since the description is based on recording the orientation of the edges of a shape, rotational transfer would not normally occur. One of the cases in which it would be expected to occur is where a shape contains some special feature which might be described separately from the remainder of the shape. The general purpose of the present experiment was to investigate transfer to rotated shapes in the goldfish and to answer the question of whether adding a feature to a shape may mediate transfer when the shape is rotated.

Different fish were trained to discriminate between a circle and one of the four shapes shown in the first row of Figure 1. The word "square" will be used to refer to these four shapes; the individual shapes will be termed normal square, normal square-k, diamond, and diamond-k where the hyphenated "k" refers to the presence of a knob. The specific aims of the experiment were as follows.



	SHAPES	GROUP NS	GROUP NS – K	GROUP D	GROUP D-K
			TRAINING SHAPES		
		■ 95	■ 95	◆ 95	◆ 92·5
			Shape+Knob rotated		
1	45°	◆ 44	◆ 70*	■ 50	■ 60
2	– 45°		◆ 55		■ 80*
3	135°		◆ 57·5		■ 47·5
4	–135°		◆ 42·5		■ 57·5
			Knob only rotated		
5	90°		■ 77·5*		◆ 80*
6	– 90°		■ 75*		◆ 82·5*
7	180°		■ 75*		◆ 77·5*
			Knob-altered		
8	0°	■ 70	■ 65*	◆ 57·5	◆ 80*
9	180°	■ 80*		◆ 92·5*	
10	45°	◆ 45	◆ 57·5	■ 62·5	■ 57·5
11	±135°	◆ 42·5		■ 42·5	
			Other shapes shown with circle		
12	▬	95*	60	62·5	65
13	▮	85*	72·5	52·5	67·5
14	/	47·5	62·5	80*	90*
15	▰	82·5*	52·5	75*	80*
16	▰	32·5	45	95*	82·5*
17	▲	55	60	95*	80*
18	▼	77·5*	62·5	62·5	75*
19	◀			77·5*	85*
20	◤	82·5*	72·5*	72·5	75*
21	◢	65	62·5	82·5*	85*
22	⬟	52·5	62·5	77·5*	82·5*
23	⬟	75	57·5	62·5	62·5
24	⬢	47·5	55	77·5*	77·5*
25	⬣	70	52·5	55	47·5
26	● or ●		● 67·5		● 70
			Further pairs of shapes		
27	▬ vs. ▮	▬ 72·5	▮ 65*		
28	▼ vs. ▲	▼ 75*	▼ 60	▲ 62·5	▲ 62·5
29	■ or ◆ vs. ● or ●		■ 57·5		◆ 72·5

FIGURE 1. Training and transfer shapes. The numbers indicate the percentage of trials on which the shape shown opposite each number was treated as equivalent to the training square by each group.

As a preliminary to the main part of the experiment we determined how readily goldfish discriminate between a circle and a normal square. This problem has been little investigated in fish, and the evidence is conflicting. Herter (1949) found that a catfish readily discriminated between a normal square and circle, but Hemmings (1965*b*) working with various tropical fish, found that this discrimination was impossible without pretraining on rectangle vs. circle, yet his animals readily learnt to discriminate between a circle and a diamond without such pretraining.

Second, it is known that, having been trained to discriminate between a circle and a normal square or diamond, octopuses show no transfer when the square is rotated through 45° (Sutherland, 1959). We tested whether goldfish also fail to transfer to rotated squares. Finally, some of the fish were trained to discriminate between a circle and a square or diamond containing a knob in order to discover whether the presence of the knob would mediate transfer when the shape was rotated.

Method

Subjects and maintenance

The subjects were 16 common goldfish (*Carassius auratus*), of from 3 to 6 in. in length. During the experiments, the fish received no food outside the experimental situation. They were kept in individual tanks, of plastic-covered angle-iron with glass sides, measuring 18 × 10 × 10 in. and filled to a depth of 7·5 in., giving a capacity of about 4·5 gal. The tanks were open to the air, and their sides and bases were covered with black paper, to minimize distraction. Each tank contained a diffuser stone, and an active-carbon and glass-wool filter hung on the outside far wall. Diffusers and filters were supplied with air from electric pumps. As the tanks were open, high evaporation made daily topping up with distilled water necessary. Both water and filter materials were changed once a month, and any debris removed from the floor of the tank each morning. The water was kept at about 24°C, and out of direct sunlight to minimize algal growth.

Experimental situation

Full details and discussion of the training situation used can be found in Mackintosh and Sutherland (1963). Each animal was trained in its own tank. At the start of each trial, the fish was retained at the near end of the tank by a black Perspex detention screen with an opening in the centre of its lower edge, closed by a sliding door. At the far end of the tank was the discrimination screen, of white Perspex, with a central vertical partition projecting at right angles down the length of the tank. The shapes to be discriminated were attached to the discrimination screen on either side of this partition; their centres were 6·5 cm. from the bottom of the screen, and midway between the partition and the tank wall. The shapes were made of ¼ in. black Perspex. At the centre of each shape and projecting towards the fish was a black feeding trough with a shallow hole drilled in its sloping front surface, to hold the bait (Hykro Concentrated Goldfish Food). The discrimination screen was 12 in. from the detention screen.

Pretraining

The fish were first trained to take food from a trough presented by itself on either side of the discrimination screen. Once a regular response had been achieved, the detention screen was introduced, and the fish were trained to go through the opening as soon as the door was raised, and take the food from the trough. Each fish received 100 trials of this pretraining; they were then taking 10 to 20 sec. to reach the trough once the door had been lifted. All subjects were then given 20 trials with the positive shape alone, presented on either side of the partition. At the end of pretraining response latencies averaged 6 sec.

Training

Immediately after pretraining, training proper was begun, with both shapes in place. The positive shape was baited with a pellet of food, the negative with a pebble resembling the food. The side occupied by the positive shape was determined by Gellermann orders.

The criterion of response was the fish's snapping at one of the shapes. If, as occasionally happened, the fish snapped at the positive shape but missed the food, the negative shape was covered and the trial continued until the food was taken, when both screens were removed immediately. At first, if the fish chose wrongly, it was allowed to correct immediately, but later this technique was modified to curb developing position habits in some of the fish. In the new procedure, after a wrong choice, a large piece of white Perspex was inserted into the tank and used to chase the fish back behind the detention screen, and the trial was re-run until the fish chose correctly. After three successive mistakes within a trial, the fish was forced to the correct side by inserting a transparent Perspex screen in front of the negative shape.

Control and transfer tests

At the end of training, control tests, in which both shapes were baited with food, were alternated with further training trials, to make sure that the fish had not learned to discriminate in terms of any perceived difference between the pebble and the food. No correction was given after an incorrect choice during control tests. In the final stage of the experiment training trials were alternated with transfer tests, in which new shapes were shown. In transfer tests, both shapes were baited with food, and no correction was given.

Experimental design

There were four experimental groups, each of four fish. Each group was trained to discriminate between a circle and a square of equal area. Different groups were trained with different square shapes, as shown in row 1 of Fig. 1. All shapes were 17·5 cm. in area. Two fish in each group were trained with circle positive, two with circle negative. Ten trials a day were given with an intertrial interval of about 20 min. The immediate correction technique was used for the first 100 trials, then the modified method was substituted. The subjects received 140 training trials, followed by one day of alternating training and control trials (five trials of each type). After the control tests, animals were given ten transfer tests with each of a number of pairs of transfer shapes. Five transfer tests were given each day, each preceded by a retraining trial.

The shapes used in transfer tests for each group are shown in Fig. 1. Except in tests 27, 28 and 29 the circle was always one of the two shapes shown. The ten trials on each transfer pair were completed before the next pair was presented. The transfer tests on different shapes were not given in a predetermined order since some of the later tests were inserted in order to test hypotheses suggested by the earlier tests. Two considerations suggest that animals' performance on transfer tests was not influenced by the serial order of the tests. First, performance on the retraining trials remained constant throughout the whole series. Second, as a control against order effects, we repeated the first transfer test (test 1) at the end of the series. No significant change in scores occurred: the amount of change for each group between the two occasions on which this test was presented ranged between 2·5 and 10 per cent and two groups showed more transfer, two less.

Results

Initial learning

All goldfish had met a criterion of 18 or more correct responses over two days by the end of training. There were no significant differences between the scores of the four main groups either in terms of trials to meet criterion or in terms of percentage correct; the average number of days to criterion for each of groups NS, NS-K, D, D-K was 10, 7, 7, 10, respectively, and the average number of correct responses made per group over the first five days training ranged between 73 and 78 per cent. During the retraining trials given on day 15, the fish averaged 94 per cent correct as compared with 89 per cent correct on the interspersed control trials; since the difference was not significant, this was taken as evidence that the fish had not solved the discrimination merely in terms of the difference between the food pellet and the pebble with which the shapes were baited in training trials.

Over the first three days of training, animals trained with square positive scored more highly (68 per cent correct) than animals trained with circle positive (53 per cent correct): this was true within each of the four groups and the overall difference between the scores of animals trained in opposite directions was highly significant ($P < 0{\cdot}001$ by a Wilcoxon test for unpaired replicates). The effect had largely disappeared by the end of training and was no longer significant over the last three days of training, during which period animals trained with circle positive averaged 86 per cent while animals trained with square positive averaged 90 per cent correct. There was therefore at the beginning of training a strong preference for the squares over the circle. This preference may have been the result of the pretraining to approach the trough, which was of square cross-section.

Transfer tests

The detailed results of the transfer tests are shown in Figure 1. The first row shows performance on retraining trials and it will be seen that the original discrimination was maintained at a high level throughout the transfer tests. The remaining numbers in the figure show the per cent of transfer trials on which the shape drawn opposite each number was treated as the original square by each group. The scores are an average of the number of times the two animals trained with square positive took the transfer shape drawn and the number of times the two animals trained with square negative took the other member of the transfer pair. Except in tests 27, 28 and 29 the other member of the pair was always the original circle.

Since there were only four animals in each group, the statistical treatment of the results presents some difficulties. A correct response on a transfer test occurs when a subject trained with circle positive approaches the circle and when a subject trained with circle negative approaches the transfer shape. On the hypothesis that the circle and a given transfer shape are not being discriminated from one another, the probability of all four animals making more than 50 per cent correct responses is 0·062 by a one-tailed binomial test, and such scores are marked with a star in Figure 1: it is legitimate to use a one-tailed test since (except in tests 27, 28 and 29) direction of transfer can be predicted from the presence of the circle. The significance of differences in performance on transfer tests within a group was evaluated by the signs test—within one group of four subjects the value of P can again not be lower than 0·06, though in several instances it was possible to do this statistic on more than four animals by combining groups. Finally, the significance of differences between the scores of different groups was evaluated by the Mann-Whitney U test.

The validity of the transfer test scores is attested by their internal consistency. Thus, in tests 12 and 13 a horizontal and vertical rectangle were presented with the original circle: group NS showed more transfer in these tests from the square to the horizontal than to the vertical rectangle and vice versa for group NS-K ($P = 0{\cdot}029$

for significance of difference between groups). In test 27 a horizontal and vertical rectangle were presented together and group NS again showed transfer from the original square to the horizontal, whereas group NS-K transferred from the original square to the vertical rectangle ($P = 0{\cdot}014$ for significance of difference). In tests 17, 18, 20 and 21, groups NS and NS-K treat the triangle with a horizontal line at the top as more like the original square than the triangle with an apex at the top ($P < 0{\cdot}008$ based on eight subjects) whereas the opposite is true for groups D and D-K ($P < 0{\cdot}008$). In these four tests the triangles were presented separately each paired with a circle but exactly the same result is obtained when an upright and inverted triangle are presented simultaneously (test 28, $P < 0{\cdot}003$, based on 16 subjects).

The detailed results of the transfer tests support the following conclusions.

(i) Animals learned the orientation of the contours of the square shapes. There was no transfer from normal square to diamond nor from diamond to normal square (test 1); the eight subjects trained on the normal squares transferred to horizontal and vertical rectangles ($P < 0{\cdot}004$ for all eight subjects) but not to oblique rectangles whereas subjects trained on the diamonds transferred to oblique but not to horizontal and vertical rectangles (tests 12, 13 and 14, $P < 0{\cdot}001$ for significance of difference between groups).

It seems likely that group D learned not only the oblique orientation of the lines of the diamond but also learned that there was an angle at the top of the shape: this group shows better transfer to a parallelogram with point at the top (test 16) than to a parallelogram with a straight line at the top (test 15, $P = 0{\cdot}062$ for significance of difference) despite the fact that both figures contain the same amount of oblique contour. Some of the results quoted under (ii) also support this conclusion.

(ii) The animals' behaviour was controlled by the tops of the shapes much more than by the bottom halves. In tests 17 to 21, there was better transfer from normal square to triangles with a straight line at the top and a point at the bottom than to triangles with a point at top and a straight line at bottom ($P = 0{\cdot}016$ for significance of difference) whereas the reverse is true for animals trained on a diamond ($P = 0{\cdot}016$). Transfer to pentagons and hexagons is also determined by whether there is a straight line or two oblique lines meeting at a point at the top (tests 22 to 25). Once again, animals in groups NS and NS-K showed better transfer when the transfer shape had a straight line at the top, whereas groups D and D-K show better transfer when there is a point at the top. Animals trained with the normal square showed more transfer from the square to a horizontal rectangle containing a long horizontal line at top than to a vertical rectangle with only a short horizontal line at the top (tests 12, 13 and 27).

(iii) Animals trained with knob shapes did detect the presence of the knob since transfer falls off if it is omitted (tests 8 and 10); also there was some transfer from the square to a circle with a knob presented with the original circle (test 26). Moreover, when a knob is added to the square for animals trained without the knob, performance again falls off and the drop in transfer is greater when the knob appears in the top half of the shape (test 8) than when it appears in the bottom half (test 9, $P < 0{\cdot}003$ for significance of difference): this result would be expected from (ii) above.

(iv) The presence of the knob is sufficient to mediate transfer from normal square-k to diamond-k and vice versa provided the knob appears in the upper half of the shape (tests 1 and 2). The knob does not mediate transfer when it is rotated through 135° so that it appears in the bottom half of the shape (tests 3 and 4). It cannot be determined whether this is because animals had learned the original position of the knob or whether they simply fail to notice it when it appears in the bottom half of the shape. Although the presence of the knob was undoubtedly detected by groups trained with it present, animals' performance was determined more by whether the main contours of the shape were horizontal and vertical or oblique than by whether or not the knob was present, particularly in the case of animals trained on diamond-k. Thus, when the training square with knob removed is presented with a circle with a knob added in its original position, there is slight transfer from the original square shape to the square without the knob (test 29); there is also transfer to the original shape with knob removed when presented with a circle and again this transfer is better for group D-K than for group NS-K (test 8).

Discussion

Theory

Sutherland (1968), following Clowes (1967), has outlined a general theory of visual shape recognition in animals: it will be sufficient to state the theory briefly here without attempting to justify it and then to see how the present results bear upon the theory.

Two mechanisms are involved in shape recognition—a processor and a store. The processor detects and labels local features present in the input shape: for example, a given unit may fire to a horizontal edge present in one part of the input array, another to a break in the continuity of a line at a particular point and so on. The processor preserves information about the spatial relationships of the different features to one another since the retina is mapped topologically onto each layer of feature detectors. When a visual stimulus is presented and attended to, one of two things happen. Either a description of the feature detectors firing is written into the store or the pattern of feature detectors firing is matched to an existing description in the store. The former process corresponds to learning, the latter to recognition. Sutherland (1968) and Sutherland and Williams (1969) have argued that the stored description of a shape is highly abstract. When a description of a shape is stored, the form of the description will determine what further shapes are classified as the same or different from the original shape.

Since a square will be described as a concatenation of horizontal and vertical lines and a diamond as a particular configuration of oblique contours, no transfer should occur between these two figures. If, however, the presence of a knob is separately encoded so that the stored description of a square bearing a knob refers to the presence of the knob, then we might expect to find transfer between normal square-k and diamond-k. In the light of the evidence from the present experiment we now consider the type of description that goldfish store about the different squares.

(i) Animals trained with the normal square seem to have stored a rule describing it almost entirely in terms of the straight horizontal edge at the top. Provided this feature was retained, transfer from the square occurred. Since, however, transfer was reduced when a horizontal line was present at the top but the sides were no longer vertical, the fish presumably also stored information about the presence of vertical sides.

(ii) Animals trained on the diamond stored information about the oblique contours at the top of the shape. They also recorded in store the presence of a point at the top of the shape (cf. above under Results). This suggests that the angle formed by the intersection of a horizontal and vertical line is stored under a different description from that used for the angles formed by the intersection of two lines in any other orientation. This is probably true of man as well as goldfish. To human beings a square appears subjectively to be a much less pointed shape than a diamond yet geometrically both shapes contain four right angles: the diamond has points facing north, south, east and west, the square has points facing north-west, north-east, south-west and south-east.

(iii) The description formed of each shape was bound by the shape's orientation—there was no generalization from normal square to diamond or vice versa. Since the presence of the knob did mediate transfer to 45° rotations, this means that the knob must have been coded as a separate feature. Unfortunately, we cannot say with certainty why goldfish do not transfer when the knob shapes are rotated so that the knob appears in the lower half; either they simply do not notice the knob in this position or they store a descriptive rule that specifies in which region of the original shape the knob appears. For example, the diamond-k could be described as "shape with oblique contours and point at top containing a knob towards the top". Since the square-k can be brought partly under the same description some transfer to it would be effected. Nor have we enough information to know how the knob itself was specified. It could have been described as a projection, but it is possible that it is simply described as a break in the continuity of a line. If the latter were true then we should expect that the presence of an indentation in the top half of the shape would also have mediated transfer for animals trained on normal square-k and diamond-k. Further work is in progress to resolve this point.

(iv) In the present experimental situation goldfish store much more information about the tops of shapes than about their bottom halves. This occurs despite the fact that the fish tend to approach the shapes from below. It seems possible that the propensity to record primarily a description of the top of the shapes occurs in other experimental situations and in other species of fish. Thus Hemmings (1965*a*) obtained the same result in several species of tropical fishes both when the discriminanda were targets against which the subject had to push its nose and when they were holes of a particular shape through which the fish had to learn to swim.

This work is part of a project on "Stimulus Analysing Mechanisms" supported by the American Office of Naval Research (Contract N62558-4791). Miss Bowman is supported by a Science Research Council research studentship.

References

Clowes, M. B. (1967). Perception, picture processing and computers. In Collins, N. L. & Michie, D. (eds.), *Machine Intelligence*, vol. 1, pp. 181–97. Edinburgh: Oliver & Boyd.

Deutsch, J. A. (1955). A theory of shape recognition. *Br. J. Psychol.* **46**, 30–7.

Hemmings, G. (1956*a*). Pretraining and transfer in shape discrimination. Unpublished Ph.D. thesis: University of Hull.

Hemmings, G. (1956*b*) The effect of pretraining in the circle/square discrimination situation. *Anim. Behav.* **13**, 212–15.

Herter, K. (1949). Zur Psychologie und Sinnesphysiologie der Zwergwelse (*Ameiurus nebulosus* (LSR)). *Biol. Zbl.* **68**, 77–95.

Mackintosh, J. and Sutherland, N. S. (1963). Visual discrimination by goldfish: the orientation of rectangles. *Anim. Behav.* **11**, 135–41.

Sutherland, N. S. (1959). Visual discrimination of shape by Octopus. Circles and squares, and circles and triangles. *Q. Jl exp. Psychol.* **11**, 24–32.

Sutherland, N. S. (1963). Shape discrimination and receptive fields. *Nature, Lond.* **197**, 118–22.

Sutherland, N. S. (1968). Outlines of a theory of visual pattern recognition in animals and man. *Proc. R. Soc.* B, **171**, 297–317.

Sutherland, N. S. and Williams, Clare (1969). Discrimination of checkerboard patterns by rats. *Q. Jl exp. Psychol.* **21**, 77-84.

Received 1 *August* 1969

Q. Jl exp. Psychol. (1969) **21**, 339–345

PROACTIVE INTERFERENCE AND SPONTANEOUS ALTERNATION IN RATS

A. W. STILL†

University of Cambridge

Two pairs of trials were given to naive rats in a T-maze. The first trial of a pair was forced to one arm, the second was free, and the delay between trials was 25 min. or 10 hr. Rats left in the maze for 15 sec. following the forced choice showed increased alternation on the second pair of trials at the 25-min. delay as compared with the 10-hr. delay. This was interpreted as due to a decline in alternation with delay on this pair of trials, together with an overall increase in alternation between the first and second pairs, and is consistent with the proactive interference theory of forgetting. Rats left in for 2 min. showed a decline in alternation with delay on the first pair of trials, but not on the second; this is inconsistent with the proactive interference theory.

In this investigation rats are used to test the proactive interference theory of forgetting (PI theory), according to which forgetting is due to interference from the traces of events or learning that occurred prior to the materials to be remembered. A strong form of the theory would be that PI is a *necessary* condition for forgetting, while a weak form would be that it increases the rate of, but is not necessary for, forgetting. As Brown (1959) has pointed out, PI cannot on its own be a *sufficient* condition; this is because forgetting must be due to some process occurring after the material to be remembered, and therefore cannot be due to an increasing number of proactively interfering events; there must be either an increasing number of retroactively interfering events (events occurring after the event to be remembered), or a change in the traces of the proactively interfering events, or a change in the traces of the events to be remembered; but the hypothesis of trace change is additional to the PI theory.

Underwood (1957) and Underwood and Postman (1960) present evidence favouring the strong form of the PI theory. Underwood (1957) shows that the more experienced a subject is in the learning of nonsense syllables, the more forgetting there is over 24 hr., and he explains this as being due to the increasing amount of proactively interfering material as the number of lists already learned is increased. Furthermore, the small amount of forgetting found in naive subjects can probably be accounted for without bringing in traditional explanations of forgetting such as trace decay and retroactive interference. In order to give a complete explanation of forgetting in experienced subjects, Underwood and Postman (1960) suggest that, to learn a new list of paired associates, interfering associations still present from previous lists must first be extinguished; these extinguished associations will recover spontaneously with time, and will therefore regain their power to interfere with the material to be remembered.

---

† Present address: Department of Psychology, Durham.

Evidence against this form of the PI theory has been presented by Warr (1964) and Gleitman and Steinman (1963). Warr showed that increased forgetting among experienced subjects could be partly explained as a function of their more rapid learning compared with naive subjects; while Gleitman and Steinman, using rats, failed to find the recovery of an extinguished habit predicted by the Underwood and Postman theory. A more direct test of the strong form of the PI theory could be made, if only it were possible to use completely naive subjects, who would be predicted to show no forgetting at all; but, as Underwood recognized, such ideal naivete seems impossible to achieve with humans, where past experience is difficult to control, so that one cannot be certain that the subject has not come across material that is similar to the material to be remembered and therefore interfering. Past experience is easier to control with animals, but if animals are trained on some task, incorrect habits or hypotheses may be incidentally acquired during training before learning has taken place, and may constitute proactively interfering material. But a form of memory may be tested *without* training by making use of the tendency in rats to alternate spontaneously in a T-maze (Dember and Fowler, 1958).

In order to be able to alternate their choice of alleys in a T-maze, rats must retain information as to which was the last alley entered, and this must be information about the stimuli encountered, rather than the responses made (Montgomery, 1952; Glanzer, 1953). Furthermore, alternation occurs even with delays between trials of several hours (Walker, 1956; Zeaman and House, 1951), so that the information must be stored in the form of a memory trace that is more than the transient stimulus satiation suggested by Glanzer (1953*b*). Douglas (1966) presents evidence suggesting an alternative possibility, namely that a rat avoids the odour trail that it laid down on the previous trial. However, Still and Macmillan (1969) show that, while the presence of an odour trail in a maze may be necessary for alternation to take place, it is not used as a cue guiding choice, but at most as a cue for familiarity, indicating, perhaps, that the maze on trial 2 is the same as on trial 1. Thus it makes no difference whether the odour trial leads down the arm previously chosen, is present in the stem only, or is reversed so that it leads from the arm *not* previously chosen into the stem; in the latter case the rat enters the arm with an odour trail in it.

If the information is stored in the form of a memory trace it follows that rate of alternation may be taken as an index of memory, and in particular that the decline in alternation with delay between trials will represent a kind of forgetting function. Using this index, a direct test of the strong form of the PI theory may be made by giving a pair of trials in a T-maze to rats that have not had any previous experience of maze arms; if PI is a necessary condition of forgetting, alternation should not decline as a function of delay between trials. However, if the rats are given a further pair of trials, there will be PI from the first pair of trials, and alternation should decline with delay. In the experiments described below these tests of the PI theory are made by comparing a 25-min. with a 10-hr. delay. These particular delays were chosen because (a) it is between them that the rate of decline of alternation with delay appears to reach a maximum in experienced rats (Walker, 1956; Zeaman and House, 1951), and (b) a minimum delay of as long as 25 min.

was felt to be necessary to eliminate the effect of the nervousness shown by rats, particularly naive ones, when trials are given in quick succession. In the first experiment the brightness difference between the maze arms was varied in order to test a hypothesis put forward by Dember and Fowler (1958) to explain why Montgomery (1951), unlike most experimenters on spontaneous alternation, found that alternation declined rapidly as a function of delay between trials; choice behaviour reached chance level after only about one minute. Dember and Fowler predict that alternation will decline gradually with delay over several hours, only if the maze arms are visually distinct; if the brightness, etc., are identical, the decline should be complete within a minute, so that alternation should not occur at either of the delays of these experiments.

The predictions from the strong form of the PI hypothesis are complicated by evidence that, for reasons that seem to have nothing to do with PI, the general level of spontaneous alternation tends to be lower in naive than experienced rats (Montgomery, 1952; Walker, 1956). However, even if this is so, the PI theory still predicts that only experienced rats will show a decline in alternation with delay.

Experiment I

Method

Subjects

Eighty experimentally naive female hooded rats, obtained from N.I.M.R., Mill Hill, and approximately 4 months old.

Apparatus

An aluminium T-maze with both arms, and the stem, 24 in. long. The alleys were 4 in. wide, 5 in. deep, and covered by a clear Perspex roof. Either arm or the stem could be closed at the choice point by means of guillotine doors. The walls of the left arm were painted grey, while those of the right arm were grey for one group, black for the other; otherwise the maze was unpainted. Lighting was provided by two 30 W strip-lights 18 in. above, running parallel to, the maze arms.

Procedure

The rats were satiated for food and water. On each of the 8 days prior to the experiment, they were put onto a rectangular table top, 18 × 60 in., in groups of 11 or 12, and handled and allowed to explore for 5 min. After the 8th day they were randomly divided into four groups of 20 rats each; for two groups both maze arms were grey (GG), while for the other two one arm was grey, one black (GB); one of the GG and one of the GB groups were given 25 min. between trials, the others were given 10 hr. The four groups were thus: GG/25 min., GG/10 hr., GB/25 min. and GB/10 hr. On the ninth to twelfth days the rats were tested, the GG groups on the ninth and eleventh, and the GB groups on the tenth and twelfth days; the procedure on these 2 days of testing was as follows:

Day 1

One pair of trials was given, the first forced by means of a guillotine door blocking access to one arm, the second free, with the appropriate delay, 25 min. or 10 hr., between the trials. The 10-hr. group were given their first trial in the morning, the second in the evening; half the 25-min. group were run in the morning, half in the evening. Half of each group were forced to one arm, half to the other. The rats were put into the maze facing the choice point. They were left in the maze for 15 sec. following the forced choice; following the free choice they were left in until both maze arms had been entered and were removed 15 sec. after entering the arm not originally chosen; the guillotine doors were not closed behind the rats following a choice. During the delay the rats were kept in their home cages in a rack situated behind the maze and covered with an opaque curtain. Those failing to choose within 5 min. were discarded. On the free trial the experimenter did not look at the tail markings (used to indicate whether the forced trial had been to the left or right) until the rat had made its choice. This meant that the free trials could not be given in a set order within cages, and the delay intervals were therefore not necessarily the same for each rat of a group; however, the mean delays were 25 min. and 10 hr., with actual delays ranging from 15 to 35 min., and 9·5 to 10·5 hr. (The difference in range of actual delays between the 25-min. and 10-hr. groups arose because the 10-hr. group were given their trials in blocks of 18 rats, the 25-min. group in blocks of 6.)

Day 2

The procedure was the same as on day 1. Each rat had the same delay between trials as on day 1; approximately half in each subgroup were forced to the same arm as on day 1; half were forced to the opposite arm. Overall, 36 rats were forced to the same arm, 33 to the opposite.

Results

Whether a rat was forced to the same arm on the 2 days, or to a different arm, did not significantly affect rate of alternation. Therefore, results from these two conditions are combined, and are shown in Table I. Ten rats were discarded for failing to choose within 5 min.; four of these were from GG/25 min., so that N for this group is 16, three from GB/10 hr. ($N = 17$), and two from each of the other two groups ($N = 18$).

TABLE I

Percentage alternation as a function of delay, experience and degree of difference between the two maze arms

Delay	GG vs GB	Day 1	Day 2	N
25 min.	GG	75	81	16
	GB	61	94	18
10 hr.	GG	67	67	18
	GB	76	76	17

As predicted by Dember and Fowler's hypothesis, there is a greater decline in alternation with delay under the GG than the GB condition, anyhow on day 1. However, none of the possible GG vs GB comparisons, either for the naive or the experienced rats (day 1 or 2) are significant by the χ^2 test ($P > 0{\cdot}1$ in all cases). Furthermore, if days 1 and 2 are combined, all groups alternate significantly above chance ($P < 0{\cdot}05$ in all cases), whereas Dember and Fowler's hypothesis would predict that alternation would only be above chance in the GG groups when the delay between trials is less than about 1 min.

To bring out any differences between naive and experienced rats in alternation as a function of delay, the GG and GB conditions were combined. There is an increase in alternation between days 1 and 2 at the 25 min. delay (by McNemar's test for the significance of changes, one-tailed $P < 0{\cdot}05$), no change at the 10-hr.

delay. A possible interpretation of these results is simply that some extraneous factor has increased alternation for the 25-min. group only on day 2. However, it is difficult to see what this factor could be, particularly as the groups were not run at clearly separate times, the 25-min. group being run at various times throughout the day. An alternative interpretation, following Walker's (1956) finding of an overall increase in alternation between days 1 and 2, is that the alternation of the 25-min. group reflects this increase, while for the 10-hr. group the increase is offset by proactive interference. Thus, as predicted from the strong form of the PI theory, there is a decline with delay on day 2 (though it is not significant), but no decline on day 1.

Experiment II

An attempt is made to replicate the findings of Experiment 1, but, instead of giving the rats 15 sec. in the maze following the first choice, half were given 2 min. and half 15 sec.

Method

Subjects

One hundred and sixty naive female hooded rats, obtained from N.I.M.R., Mill Hill, and approximately 4 months old.

Apparatus

The same as in Experiment 1 except that, (a) the roofs of the maze were of Ivorene, a substance opaque to all but diffuse light; (b) the two strip-lights were 3 in. above, and fixed to, the roofs; and (c) the left arm always had vertically striped walls, and there were stripes of PVC tape on the Ivorene roof, so that the light filtered through in stripes, while the right arm had black walls, and a plain Ivorene roof. The rats could be observed making their choices through slits between the roofs of the arms and the roof covering the choice point.

Procedure

The same as in Experiment 1 except that, (a) the difference between the maze arms was the same for all rats; and (b) on day 1 half the rats were given 15 sec. in the maze following the forced trial, half were given 2 min. The following labels are therefore descriptive of the 4 groups: 15 sec./10 hr., 2 min./10 hr., 15 sec./25 min., and 2 min./25 min. On day 2 half the rats previously given 15 sec. were given 2 min. in the maze following the forced trial, half were given 15 sec. again; similarly, half those previously given 2 min. were given 15 sec., half were given 2 min.

Results and Discussion

Rate of alternation on day 2 was not significantly or systematically affected by the condition (15 sec. or 2 min. on forced choice, forced to left or right) given on day 1. Therefore results are combined in considering alternation on day 2.

The results are shown in Table II. Eighteen rats were discarded for failing to choose within 5 min.; four of these were from 15 sec./10 hr. ($N = 36$), seven from 2 min./10 hr. ($N = 33$), none from 15 sec./25 min. ($N = 40$), and seven from 2 min./25 min. ($N = 33$). The number of rats discarded from the groups given 2 min. after the forced choice on day 1 was significantly greater than the number discarded from the groups given 15 sec. ($\chi^2 = 5{\cdot}8$, $P < 0{\cdot}02$). Observation of the rats suggests that this difference occurred because of increased emotionality on being picked up out of the maze after 2 min. compared with 15 sec.

TABLE II

Percentage alternation as a function of delay, experience and time in the maze following the forced choice

Delay	Time in maze	Day 1		Day 2	
		% alt.	*N*	% alt.	*N*
25 min.	15 sec.	70	40	73	37
	2 min.	91	33	81	36
10 hr.	15 sec.	67	36	61	33
	2 min.	70	33	81	36

When the rats are given 15 sec. in the maze after the forced choice, as in Experiment I, there is more decline of alternation with delay on day 2 than on day 1. Combining the results from the two experiments, there is a significant decline with delay on day 2 ($\chi^2 = 2\cdot85$, one-tailed $P < 0\cdot05$), while on day 1 there is 69 per cent alternation at both delays. However, unlike in Experiment I, there is in Experiment II no significant increase in alternation at the 25-min. delay between days 1 and 2. One possible interpretation of the two experiments together is that there is, for some reason, variation in the extent to which overall alternation increases between days 1 and 2, but that, imposed upon this variation, there is a constant decrease on day 2 due to delay, as would be predicted by the strong form of the PI hypothesis. (Alternatively, as mentioned before, there could be some extraneous factor in Experiment I increasing alternation at the 25-min. delay on day 2, with the remaining variations in alternation being due to chance.) Thus the results from both experiments taken together are consistent with the prediction from the strong form of the PI hypothesis.

When the rats are given 2 min. in the maze after the forced choice the results are reversed, and there is a decline in alternation with delay on day 1 ($\chi^2 = 3\cdot44$, one-tailed $P < 0\cdot05$), but not on day 2. Furthermore, there is, if anything, a *decrease* in alternation between days 1 and 2 at the 25-min. delay, an increase at the 10-hr. delay (though the differences are not significant). Thus there is a decline in alternation with delay on day 1 even in the absence of prior interfering experiences, and therefore, unless some other, more obscure, proactively interfering experiences are postulated, and provided that spontaneous alternation is a true test of memory, the strong form of the PI theory must be rejected; proactive interference is not a necessary condition of forgetting.

The explanation of the apparent interaction between delay, time in the maze following the forced choice, and experience, must await further experimental analysis of the effect of these variables upon alternation.

I should like to thank the Medical Research Council for financial support; Mr. A. J. Watson, who supervised this research, for much helpful discussion and advice; and Professor O. L. Zangwill for his continued encouragement.

References

BROWN, J. (1959). Information, redundancy and decay of the memory trace. In *The Mechanisation of Thought Processes*, Vol. II, pp. 729–45. London: H.M.S.O.

DEMBER, W. N. and FOWLER, H. (1958). Spontaneous alternation behaviour. *Psychol. Bull.* **55**, 412–28.

DOUGLAS, R. J. (1966). Cues for spontaneous alternation. *J. comp. physiol. Psychol.* **62**, 171–83.

GLANZER, M. (1953*a*). The role of stimulus satiation in spontaneous alternation. *J. exp. Psychol.* **45**, 387–93.

GLANZER, M. (1953*b*). Stimulus satiation: an explanation of spontaneous alternation and related phenomena. *Psychol. Rev.*, **60**, 257–68.

GLEITMAN, H. and STEINMAN, F. (1963). Retention of runway performance as a function of proactive interference. *J. comp. physiol. Psychol.* **56**, 834–8.

MONTGOMERY, K. C. (1951). Spontaneous alternation as a function of time between trials and amount of work. *J. exp. Psychol.* **42**, 82–93.

MONTGOMERY, K. C. (1952). A test of two explanations of spontaneous alternation. *J. comp. physiol. Psychol.* **45**, 287–93.

STILL, A. W. (1966). Repetition and alternation in rats. *Q. Jl exp. Psychol.* **18**, 103–8.

STILL, A. W. and MACMILLAN, A. St. C. Odour trail and spontaneous alternation. Submitted for publication to *Psychon. Sci.*

UNDERWOOD, B. J. (1957). Interference and forgetting. *Psychol. Rev.* **64**, 49–60.

UNDERWOOD, B. J. and POSTMAN, L. (1960). Extraexperimental sources of interference in forgetting. *Psychol. Rev.* **67**, 73–95.

WALKER, E. L. (1956). The duration and course of the reaction decrement and the influence of reward. *J. comp. physiol. Psychol.* **49**, 167–76.

WARR, P. B. (1964). The relative importance of proactive inhibition and degree of learning in retention of paired associate items. *Br. J. Psychol.* **55**, 19–30.

ZEAMAN, D. and HOUSE, B. J. (1951). The growth and decay of reactive inhibition as measured by alternation behaviour. *J. exp. Psychol.* **41**, 177–86.

Received 1 *August* 1969

Q. Jl exp. Psychol. (1969) **21**, 346–355

WHEN IS A WEIGHT NOT ILLUSORY?

HELEN E. ROSS

Psychological Laboratory, University of Cambridge[†]

Weight illusions occur whenever some aspect of an object—such as its size, material or colour—arouses the expectation that its weight will be heavier or lighter than it actually is. The direction of the illusion normally contrasts with the expected weight. When objects are hidden from sight and lifted by strings they can provide no misleading cues, and a correct weight-expectation should be achieved after one or two trials. When a visible object has the same physical and apparent weight as a hidden object, it can be defined as non-illusory. Weighted tins and polystyrene blocks of various sizes were compared with hidden weights. Tins were found to be non-illusory when their density was about 1·7, and polystyrene blocks when their density was about 0·14.

Weight illusions may be due to a central scaling process which enables a wide range of weights to be estimated, different ranges being selected according to the expected value of the weight. If the selected range is inappropriate an illusion occurs. Changes in expected value could also allow for the operation of "weight-constancy" during changes in proprioceptive stimulation.

Introduction

If two objects are equal in weight but different in one other aspect (e.g. size, material, colour), their apparent weights will differ. In this case the difference in apparent weight will normally *contrast* with the difference that might be expected from the appearance of the objects. Illusions of *assimilation* have also been reported; but these seem to occur only when the sensory information is inadequate, and the subject relies upon the appearance of the objects (Howard, 1954; Ross, 1966). Werber and King (1962) report an abnormally high incidence of reversed illusions under normal conditions, but since the larger object was always presented second the results are probably confounded by the time-order error.

The most familiar of the weight illusions is the size-weight illusion (Charpentier, 1891), where the larger of two equal weights feels lighter. A similar illusion occurs with differences in material (Wolfe, 1898; Seashore, 1899): if a wooden container and a metal container are of equal weight and volume, the metal container feels lighter. Another example is the colour-weight illusion: darker objects are expected to be heavier than brighter objects (Bullough, 1907), but feel lighter after lifting (De Camp, 1917).

Presumably these illusions are due to learning the normal relations between the appearance of objects and their weight. They also seem to depend on a fairly high level of perceptual integration. Klüver (1933, p. 54) failed to find the size-weight illusion in monkeys: and several authors have claimed that it is reduced, absent, or sometimes reversed in very backward children (Demoor, 1898; Claparède, 1903;

[†] Present address: Department of Psychology, University of Stirling.

Rey, 1930; Koseleff, 1936; Jenkin and West, 1958). Doll (1913) claimed that the illusion occurred in backward children after a mental age of eight years, and Ohwaki (1953) after a mental age of four to five years. General brain deterioration in adults seems to have little effect (Nyssen and Bourdon, 1955). However, Halstead (1945) reported the absence of the illusion in a patient with a left frontal lesion, along with other defects of perception and organization.

The weight illusions due to size, material and colour seem to depend upon weight-cues which have been very thoroughly learned. They can be reduced by massed practice, but the effect is often only temporary. Short-term illusions can also be induced by suitable recent experience: Uznadze (1939) showed that after repeated presentation of a heavy weight to the left hand and a light weight to the right hand, the left hand of two equal weights felt lighter. Wetenkamp (1933) found that after practice with materials of abnormal densities, the material-weight illusion was reversed when the same materials were presented with their normal densities. Furthermore, these contrast illusions can be produced by purely *imaginary* practice (Natadze, 1960).

Both the long-term and the short-term illusions can be classified as "contrast with the expected weight." The changes in expectation seem to entail central rather than peripheral processes. Changes in muscle-tension are not necessarily involved, since the illusion can occur from pressure alone, without lifting (Bochorishvili, 1927; Uznadze, 1931). Changes in the pressure receptors are not essential either, since the illusion can occur when the weights are lifted by anaesthetized hands (Ross, 1966).

Since the same physical weight can give rise to a variety of apparent weights, one may well ask which of the apparent weights is "correct." Are some apparent weights more illusory than others? If illusions are in fact due to contrast with the expected weight, this implies that the apparent weight of an object is not illusory if it is correctly expected. However, this tautology does not enable one to determine the expected weight of an object experimentally. A practical approach to the problem is to display a variety of objects of different weights but the same size (or the same weight but different sizes) and ask "Which of these objects feels the right weight for its size?" I tried this experiment but found that subjects either could not understand or could not answer the question. A better operational definition would be to say that the weight of an object is expected correctly if the object is lifted with an appropriate force. Unfortunately small errors of lifting are difficult to detect.

A more fruitful approach to the problem is to use hidden weights. If the objects are hidden from sight and lifted on strings there are no visual or tactile cues to their weight before lifting. Under this condition expectations are likely to be imprecise or incorrect at first; but after one or two attempts a correct expectation should develop unhindered by misleading cues. When the objects are visible the weight-cues are very compelling, and the illusions are difficult to extinguish. This is probably why we shut our eyes when trying to compare the weights of objects or "guess the weight of the cake."

If hidden weights are non-illusory they can be used as standards by which to measure visible weights. When hidden and visible objects of the same physical weight have the same apparent weight it can be argued that the visible object has the expected combination of size and weight and is not illusory. A definition of this sort was suggested by Huang (1945*a*), but he did not make use of it experimentally. The advantage of this definition will be discussed in more detail after presentation of some data on the size-weight and material-weight illusions.

Method

Three experiments were performed, each involving 15 subjects. The 45 subjects were undergraduates and research students of both sexes. In each experiment one hidden standard weight was used, and a range of visible weights in each of four sizes. All the containers were painted grey. The weights used in the three experiments were as follows:

(1) Four sizes of expanded polystyrene blocks, with volumes of 4005, 1780, 800 and 550 cm^3. Their weight was varied by inserting lead shot in holes cut from the base of the block and then sealed. Sections of 2BA threaded rod were inserted in the tops of the blocks, as vertical handles. The hidden standard weighed 121 g. The weights in each size ranged from 90–165 g. There were 3-g. intervals from 112–130 g. and 5-g. intervals outside that range.

(2) Four sizes of tins, with volumes of 326·9, 194·4, 38·60 and 18·54 cm^3. The two larger tins were originally coffee tins (Nescafé), and the two smaller tins film containers (Kodak and Ferrania). The tins were weighted with lead shot and wax. Sections of 2BA threaded rod were fitted in the lids as handles. The hidden standard weighed 120 g. It was not convenient to produce the same range of weights in all sizes of tins, because of the different minimum weights of empty tins and maximum weights of full tins. The ranges used were 100–180, 90–170, 70–150 and 60–130 g., respectively. All weights differed by 5-g. intervals.

(3) Four sizes of tins, with volumes of 962·7, 326·9, 194·4 and 38·60 cm^3. The largest size was originally a coffee tin (Sainsbury's), and the others were as described above. The tins were prepared as before. The hidden standard weighed 210 g. The range of weights covered in all sizes was 160–270 g. The weights differed by 10-g. intervals except between 190–250 g. where there were 5-g. intervals.

All the weights quoted above include an extra 10 g. due to the attachments by which they were lifted inside a box. A diagram of the box is shown in Fig. 1.

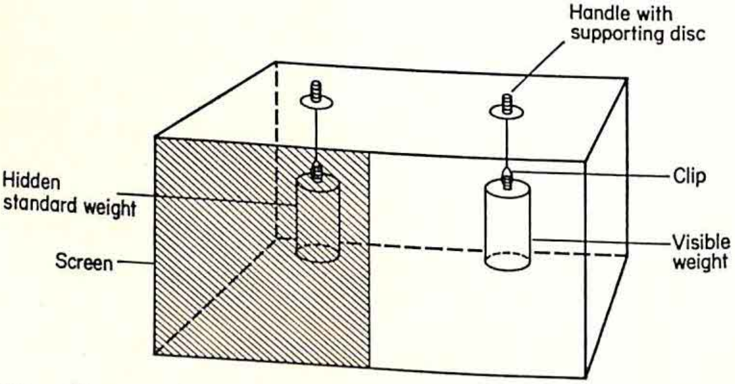


FIGURE 1. Diagram of the box used for comparing visible and hidden weights.

The box was 33 cm. high, and 43 × 29 cm. on top. The box was open at the back and front, though half the front was hidden by a screen. In the top of the box were two holes, 20 cm. apart. Through each hole hung a crocodile clip, attached to a piece of string 12 cm.

long. The top of each string was tied to a 2BA bolt, which was supported above the hole by an aluminium disc 2·5 cm. in diameter. The experimenter hung the weights inside the box, by inserting their vertical handles inside the crocodile clips. The subject lifted the weights by means of the bolts which were visible on top of the box. The standard weight was hidden behind the screen on the subject's left, but the variable weight on the right could be seen.

The testing procedure was the same for all three experiments. The subject was seated in front of the box, and the hidden standard was suspended behind the screen. This weight could be of any size since it was not seen. The four different sizes of standards were in fact used in turn, but this had no effect upon the results. The experimenter then selected the weight which was equal to the standard from one of the sizes of variable weights, and suspended it in the unscreened half of the box. The subject was asked to compare the hidden visible weights by lifting them alternately with the same hand, looking carefully at the visible weight. He lifted them as often as he liked, until he was able to report whether the visible weight felt heavier, lighter or the same as the hidden weight. If it felt heavier or lighter, he was given a correspondingly lighter or heavier weight (differing by one step interval) on the next trial. Different visible weights were presented until the subject judged that the two weights felt equal. The first judgement of "equal" was accepted as the score. The procedure was then repeated for the three other sizes of weights. The order of testing for the four different sizes was randomized between subjects. The whole test lasted between 20 and 40 min., depending upon the speed of the subject's decisions.

Results

The results for individual subjects are shown in Tables I to III. The mean weights judged equal to the hidden standard are shown graphically in Figure 2, as a

TABLE I

Weight (g.) of the different sized polystyrene blocks which were matched with a hidden standard of 121 g.

Subject	Volume of block (cm.^{3}) 4005	1780	800	550
1	145	128	118	118
2	127	124	115	112
3	127	124	118	115
4	140	124	121	115
5	140	130	121	115
6	135	121	118	112
7	127	130	118	115
8	121	115	118	110
9	165	140	140	130
10	145	127	124	128
11	140	135	127	124
12	118	115	118	105
13	130	127	121	124
14	127	130	115	105
15	145	130	127	121
Mean	135·5	126·7	121·3	116·6
Density of match	0·0338	0·0712	0·1516	0·2120

TABLE II

Weight (g.) of the different sized tins which were matched with a hidden standard of 120 g.

Subject	Volume of tin (cm.^{3}) 326·9	194·4	38·6	18·54
1	135	125	120	105
2	170	145	125	120
3	155	145	120	110
4	105	95	85	75
5	135	130	120	115
6	150	130	105	90
7	135	125	110	100
8	140	125	110	95
9	140	135	115	110
10	130	120	115	100
11	135	125	115	110
12	140	135	115	110
13	130	120	115	105
14	125	120	105	85
15	130	120	115	100
Mean	137	126·3	112·7	102
Density of match	0·4191	0·6499	2·9188	5·5016

TABLE III

Weight (g.) of the different sized tins which were matched with a hidden standard of 210 g.

Subject	Volume of tin (cm.^{3}) 962·7	326·9	194·4	38·6
1	235	230	215	220
2	240	220	210	215
3	240	225	220	215
4	200	190	180	170
5	245	230	225	205
6	215	200	195	195
7	220	220	195	170
8	230	225	220	210
9	260	240	230	220
10	210	200	190	190
11	270	260	225	220
12	230	220	215	200
13	225	215	215	215
14	250	225	205	200
15	240	225	205	205
Mean	234·0	221·7	209·7	203·3
Density of match	0·2431	0·6781	1·0785	5·2677

function of the logarithm of the density of the visible weight. The straight lines shown were fitted by an analysis of variance with a linear regression. In each of the three analyses the linear regression accounted for a significant part of the total variance, as did the differences between subjects ($P < 0{\cdot}001$).

The log densities where the hidden and visible weights were physically and subjectively equal are indicated in Figure 2. The corresponding densities were 0·1409 for polystyrene, 1·389 for the lighter set of tins and 2·015 for the heavier set. The small disagreement over the "correct" density for tins in the two experiments is not surprising considering the variability between subjects.

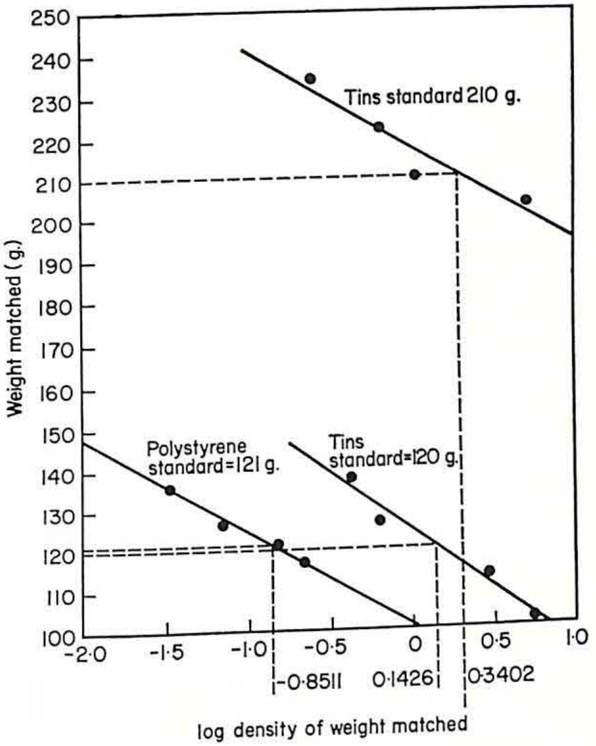


FIGURE 2. Relation between apparent weight and log density. The figure shows the log density at which a visible weight was correctly matched with a hidden standard.

Discussion

The results show that apparent weight is a linear function of the logarithm of the density (over the range used) when only the volume is changed, weight and material being constant. When the material is changed the line is shifted sideways. This can be seen by comparing the lines for polystyrene and the lighter set of tins. If these lines were not parallel they would intersect at some density where there would be no material-weight illusion. Since they are parallel (approximately), all tins and polystyrene blocks which have the same apparent and physical weight must differ by a constant log density. Similarly, if they have the same apparent weight and the same density, they must differ by a constant physical weight; in this case the tins are about 20 g. heavier than the polystyrene blocks.

In comparing the results for polystyrene and tins in this way it is assumed that similar results would have been obtained if the same group of subjects had been used for both experiments. It is also assumed that hidden standards of the same physical weight (ignoring the 1 g. difference) have the same apparent weight, regardless of whether they are compared with polystyrene or tins. Thus a point on either line has the same apparent weight as any other point on either line, since all points are equal to the same hidden standard.

The two experiments with tins cannot be compared in quite the same way, since different hidden standards were used. However, it can be seen that any point on one line differs from any point on the other line by the same apparent weight—the difference in apparent weight of the hidden standards. Since the lines are parallel, a constant difference in physical weight between tins of the same density will produce a constant difference in apparent weight, regardless of the absolute value of the density.

Other authors (Scripture, 1897; Huang, 1945*a*; Howard, 1954) have investigated the relation between density and the size-weight illusion, but their results are difficult to compare because of differences in method. Measures such as "percentage illusion" are ambiguous because of the arbitrary choice of denominator. Such studies will only become comparable when some agreement is reached on the definition and measurement of an illusion.

The importance of density can be seen in studies concerned with the relation between physical and apparent weight, when size is constant. Harper and Stevens (1948) found that apparent weight increased more rapidly than physical weight, using a fractionation method. However, the experiment was contaminated by the size-weight illusion, since the heavier weights were also denser. Warren and Warren (1956) found that if the volume was varied so that density was constant, the judged ratios between weights were fairly accurate. Baker and Dudek (1955) found a similar result for hidden weights lifted on chains.

It appears that there is a linear correspondence between physical and apparent weight, provided density is either constant or unknown. The same result should hold for different types of objects of various weights and densities, provided they are familiar and have not been "rigged." All familiar objects should be non-illusory, so their apparent weights should show the same mathematical relationships as those of hidden objects. Babington-Smith (1965, p. 149) found good rank ordering of relative weight for most ordinary objects, with a few exceptions for less familiar objects of unexpectedly high or low density (e.g. a shotgun cartridge and a papier-mâché eggcup).

What is the purpose of a scaling mechanism so liable to illusions? To say that the illusions are due to "contrast" or "set" (Uznadze, 1931, 1939; Bzhalava, 1958) may be a useful generalization, but does not explain the mechanism involved. It has sometimes been suggested that a constancy process—"density-constancy"—is responsible for the size-weight illusion (Thouless, 1931; Brunswik, 1933; Huang, 1945*a*,*b*; Howard, 1954). It was argued that we tend to perceive density rather than weight, and that this is a constancy phenomenon. However, there is no independent evidence for density-constancy apart from the size-weight illusion, and it is a semantic problem whether we choose to speak of an illusion or of constancy. The other weight illusions bear no relation to density-constancy, and the material-weight illusion gives the opposite effect. Attempts to explain the size-weight illusion by density-constancy are therefore circular, and the concept of density-constancy lacks any general explanatory power.

The term "weight-constancy" has also been used, in a variety of senses. Huang (1945*b*) used it in contrast to density-constancy. This seems an inappropriate use of the term, since constancy of perception is normally contrasted with a changing value of the proximal stimulus. In the case of weight and density, there is no change in the proximal stimulus when volume alone is changed. It would be proper to speak of density-constancy despite change of weight, but not weight-constancy despite change of density.

The term "weight-constancy" has been used in a proper sense by several authors to contrast with changing proprioceptive information, though the object lifted remains the same. Fischel (1926) and Katz (1935) used it to refer to the fact that an object feels much the same weight regardless of the manner of lifting. This applies even to amputees with artificial limbs (Katz, 1935, p. 186), and to the perception of weights lifted by mechanical apparatus (Crawford, 1961). Similarly, we can to a large extent discount changes in torque (when weights are suspended at the ends of rods) provided we have visual information (Torrey, 1963; Woodruff and Helson, 1965). We perceive the weights of objects, and not the systematic variation in stimulation due to the different positions of the arm or a mechanical apparatus.

A weight-perception system which responded to the total sensory input would have disadvantages which are probably more serious than the occasional illusions which result from the existing system. One disadvantage is that it would not allow for weight-constancy. To do this the system must be able to discount expected changes in sensory input unconnected with the object itself. Another disadvantage is that the number of distinguishable rates of nervous firing is limited, and if the available rates had to cover all possible sensory inputs discrimination between different intensities would be poor.

The perception of weight seems to be greatly influenced by the contrast between the expected and received sensory inputs, though the value of the expected input is also taken into account. This system has the advantage that a large part of the sensory capacity can be used for measuring contrast, thus enhancing contrast information at the expense of information about absolute value.

In conclusion, it is argued that the weight illusions are the result of a measuring system which takes into account the expected value of an object. This gives an opportunity for weight-constancy—the disregarding of irrelevant sensory changes—and makes maximum use of the sensory capacity for discrimination.

This work was performed while I was supported by a research studentship from the Science Research Council. I am grateful to Professor R. L. Gregory for supervising the work; to Professor R. J. Audley and Dr. D. E. Broadbent for helpful criticisms; to Miss Violet Cane for statistical help; and Professor O. L. Zangwill for encouragement and facilities.

References

Babington-Smith, B. (1965). *Laboratory Experience in Psychology*. Oxford: Pergamon Press.

Baker, K. E. and Dudek, F. J. (1955). Weight scales from ratio judgments and comparisons of existent weight scales. *J. exp. Psychol.* **50**, 293–308.

Bochorishvili, A. A. (1927). A pressure analogy to the weight illusion. *Bull. Tbilisi Univ.* **3**. (cited by Bzhalava, 1958.)

Brunswik, E. (1933). Die Zugänglichkeit von Gegenständen für die Wahrnemung und ihre quantitative Bestimmung. *Arch. ges. Psychol.* **88**, 377–418. (cited by Huang, 1945.)

Bullough, E. (1907). On the apparent heaviness of colours. *Br. J. Psychol.* **2**, 111–52.

Bzhalava, I. T. (1958). (The nature of the contrast illusion) *Vop. Psikhol*, **4**, 42–52. (Trans. in *Recent Soviet Psychology*, Ed. N. O'Connor. Oxford: Pergamon, 1961.)

Charpentier, A. (1891). Analyse expérimentale de quelques éléments de la sensation de poids. *Arch. physiol. norm. Pathol.* **3**, 122–35 (cited by Howard, 1954).

Claparède, E. (1903). L'illusion de poids chez les anormaux. *Archs. Psychol., Génève*, **2**, 22–32.

Crawford, B. M. (1961). Measures of remote manipulator feedback: Absolute judgments of weight. USAF WADD T.R. 60–591(ii) Wright-Patterson Air Force Base, Ohio.

De Camp, J. E. (1917). The influence of colour on apparent weight. *J. exp. Psychol.* **2**, 347–70.

Demoor, J. (1898). A propos d'une illusion musculaire. *J. méd. Brux.* **2**, Jan. 13th (cited by Claparéde, 1903).

Doll, E. (1913). The Demoor size-weight illusion. *Training School Bull.* **9**, 145–9 (cited by Jenkin and West, 1958).

Fischel, H. (1926). Transformationserscheinungen bei Gewichtschebungen. *Z. Psychol.* **98**, 342–65.

Halstead, W. C. (1945). Brain injuries and the higher levels of consciousness. *Res. Publs. Ass. nerv. ment. Dis.* **24**, 480–506. In *Trauma of the Central Nervous System*. Baltimore: Williams and Wilkins.

Harper, R. S. and Stevens, S. S. (1948). A psychophysical scale of weight and a formula for its derivation. *Am. J. Psychol.* **61**, 343–51.

Howard, I. P. (1954). An experimental and theoretical analysis of the size-weight illusion. *Durham Res. Rev.* **1**(5), 47–53.

Huang, I. (1945*a*). The size-weight illusion in relation to the perceptual constancies. *J. gen. Psychol.* **33**, 43–63.

Huang, I. (1945*b*). The size-weight illusion and the "weight-density illusion". *J. gen. Psychol.* **33**, 65–84.

Jenkin, N. and West, N. (1958). Perception in organic mental defectives: An exploratory study. 1. The size-weight illusion. *Training School Bull.* **55**, 5–10.

Katz, D. (1935). *The World of Colour*. London: Kegan Paul.

Klüver, H. (1933). *Behavior Mechanisms in Monkeys*. Chicago: Chicago University Press.

Koseleff, P. (1936). Eine modifikation des Charpentier-Effektes. *Psychol. Forsch.* **31**, 142–5.

Natadze, R. (1960). Emergence of set on the basis of imaginal situations. *Br. J. Psychol.* **51**, 237–45.

Nyssen, R. and Bourdon, J. (1955). Contribution to the study of the size-weight illusion by the method of P. Koseleff. *Acta Psychol.* **11**, 467–74.

Ohwaki, S. (1953). On weight perception, especially the formation of Charpentier's illusion in children; a developmental study of perception. 1. *Jap. J. Psychol.* **24**, 193–202 (*Psychol. Abstr.* 1955, **29**, 3422).

Rey, A. (1930). Contributions à l'étude de l'illusion de poids chez les anormaux. *Archs. Psychol., Génève* **22**, 285–97.

ROSS, H. E. (1966). Sensory information necessary for the size-weight illusion. *Nature, Lond.*, **212**, 650.

SCRIPTURE, E. W. (1897). *The New Psychology*. London: Scribner (cited by Huang, 1945*a*).

SEASHORE, C. E. (1899). Some psychological statistics. 2. The material weight illusion. *Univ. Iowa Stud. Psychol.* **2**, 36–46.

THOULESS, R. H. (1931). Phenomenal regression to the real object. 2. *Br. J. Psychol.* **22**, 1–30.

TORREY, C. C. (1963). *The distance-weight effect: an exploratory study in weight perception.* Unpublished dissertation. Cornell Univ. (*Dissert. Abstr.* 1964, **24**, 3858).

UZNADZE, D. (1931). Über die gewichtstäuschung und ihre Analoga. *Psychol. Forsch.* **14**, 366–79.

UZNADZE, D. (1939). Untersuchungen zur Psychologie der Einstellung. *Acta Psychol.* **4**, 323–60.

WARREN, R. M. and WARREN, R. P. (1956). Effect of the relative volume of standard and comparison object on half-heaviness judgments. *Am. J. Psychol.* **69**, 640–3.

WERBER, M. and KING, D. J. (1962). An investigation of the response mechanism of the size-weight illusion. *J. gen. Psychol.* **66**, 85–100.

WETENKAMP, L. (1933). Über die Materialtäuschung. Ein Beitrag zur Lehre von der Objektion. *Z. Psychol.* **130**, 172–234.

WOLFE, H. K. (1898). Some effects of size on judgments of weight. *Psychol. Rev.* **5**, 25–54.

WOODRUFF, B. and HELSON, H. (1965). Torque: a new dimension in tactile-kinesthetic sensitivity. *Am. J. Psychol.* **78**, 271–7.

Received 30 *June* 1968

Q. Jl exp. Psychol. (1969) **21**, 356–366

AN EXPERIMENTAL STUDY OF ANGULAR SUBTENSION

GERALD H. FISHER

Department of Psychology, University of Newcastle upon Tyne

The confusion arising over judgements of the sizes of angular forms is described. Three experiments are reported. In the first, the apparent-sizes of acute and obtuse angles are estimated. In the second, judgements of the sizes assumed by upward- and downward- and left- and right-facing angles are compared. The third experiment is concerned with tactile reproduction of angles of specified sizes. In discussion an attempt is made to clarify the logical and empirical problems involved.

Introduction

In *The Perceptions of Vision* Helmholtz stated "*. . . that as a general thing acute angles, composed of smaller angles distinctly marked, seem to be relatively too large when they are compared with obtuse or right angles that are not thus divided . . .*" (1962, p. 196). Reaffirming this, Wundt observed "*. . . that acute angles are overestimated, obtuse angles underestimated, and that the direction of the lines forming the angle varies correspondingly . . .*" (1902, p. 137). As far as we have been able to determine, both these statements were derived from the characteristic appearances of the angle illusions since neither of these authors appear to have made any independent experimental studies of apparent angular subtension. Rather, they based their statements upon the view of Hering (1861) who developed the principles of angular distortion from the evidence that, in general, interrupted distances are overestimated in size while otherwise similar uninterrupted distances are underestimated. Direct experimental study of the apparent-sizes of angles awaited Jastrow who, in introducing his widely-quoted experiment, remarked "*. . . a number of writers have based their explanations of important optical illusions upon the view (apparently not tested by experiment) that acute angles are underestimated and obtuse angles overestimated . . .*" (1892, p. 214).

Since they will be referred to throughout this study, the results of Jastrow's experiment have been reproduced in Figure 1. They indicate that angles of 30° through 75°, along with those of 150° and 165°, are *underestimated* in size; angles of 15°, in addition to those of 105° and 120°, are *overestimated*; and the angles of 90° and 135° are judged most accurately of all. These results led Jastrow to observe "*. . . It will be seen that the overestimations and underestimations can hardly be said to follow any simply formulatable law such as underestimation of acute and the exaggeration of obtuse angles . . .*" (1892, p. 215).

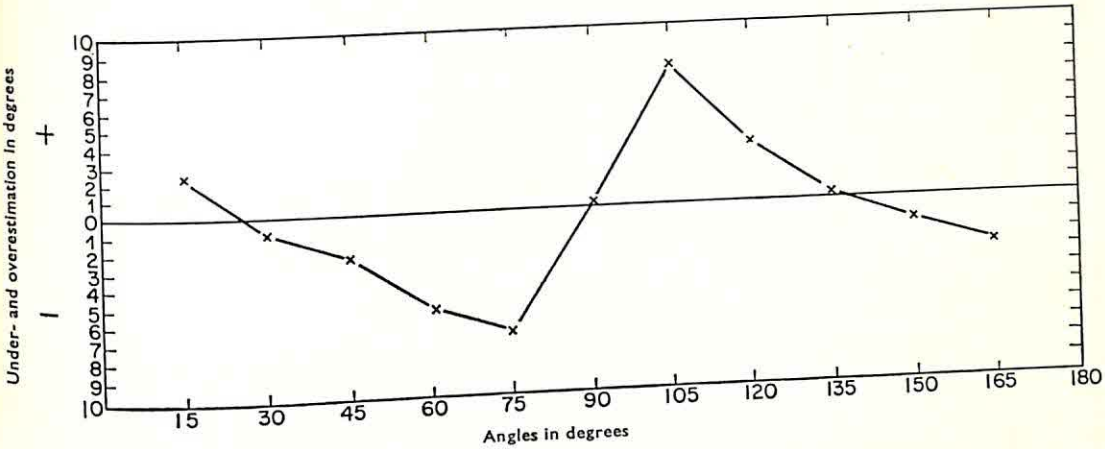


FIGURE 1. Jastrow's data referring to judgements of angular subtension. (Reproduced from *Am. J. Psychol.* (1892) 5, with permission.)

The above statements describing apparent-distortion of angular size are in direct contradiction. Helmholtz, Wundt and many others considered that acute angles are *overestimated* and obtuse angles *underestimated*. Jastrow did not specify the names of the "*number of writers*" whom he claimed found that acute angles are *underestimated* and obtuse angles *overestimated*. His own experimental enquiry revealed a wide area of *underestimation* in the case of acute angles and another of *overestimation* for obtuse angles. This outcome was largely in accordance with his original statement. A large number of studies of apparent angular subtension have been reported. To my knowledge, however, attention has not previously been drawn to this contradiction. Since apparent-distortion of angles is extremely relevant to understanding fundamental properties of space and shape perception, it is timely that an attempt should be made to clarify this problem along with certain important issues attendant upon it.

Experiment 1

Materials and procedure

Naturally, it is possible to prepare an infinite number of combinations of two lines intersecting in such a way as to define angular forms. To keep this study within realistic bounds, however, the stimulus-materials used in this experiment were restricted to the angles subtended by a horizontal straight line and a series of further lines drawn so as to irradiate through the first and fourth angular quadratures. In this way a set of 36 different angles ranging in size from 5° through 180° was defined; the increment between adjacent figures being 5°. All the angles were drawn in black line upon a white background, the boundaries of which were defined by a black circular mask. They were photographed and prepared as 2 in. × by 2 in. slides. These were exposed by rear projection on a circular shaped screen made from opal glass to a sample of 132 individual subjects. The subjects were seated in a chair constructed so as to hold their heads and bodies firmly in a vertical orientation. The amount of extraneous light in the laboratory was reduced to a minimum. The slides were rearranged in a different random order each time the experiment was carried out. All subjects were asked to make absolute judgements, in degrees, of the sizes of the angles presented and they made one such judgement only for each angle included.

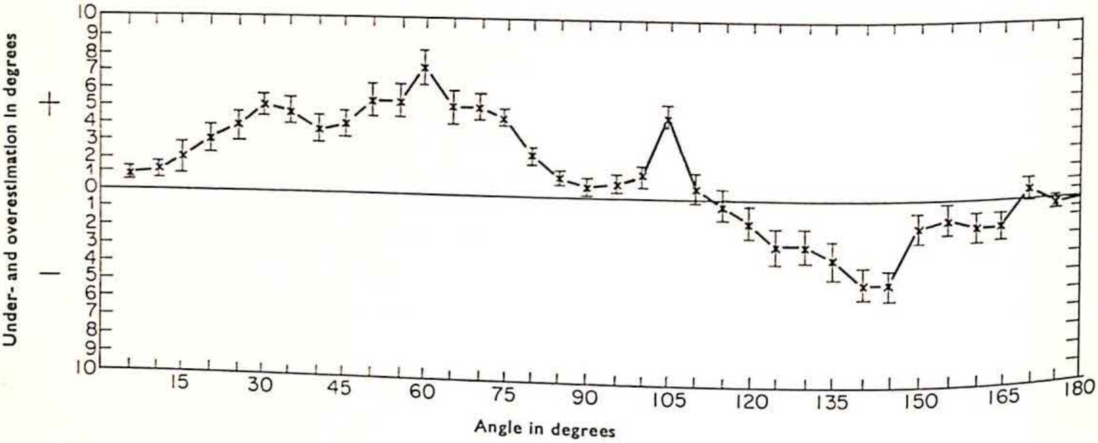


FIGURE 2. Absolute judgements of the sizes subtended by first and fourth quadrant angles.

Results

For clarity and to ensure ease of comparison with other data presented in this paper, the results of this study are shown in graphical form in Figure 2. The graph-line is defined by the mean judgements made by all subjects for each angle included. Angular size is indicated on the abscissa and the degrees of over- and underestimation are referred to on the ordinate. The range of one standard error of estimate above and below each mean value is indicated by a bounded vertical bar. These results show that there is a wide range of distortion in absolute judgements of angular subtension. It ranges from an overestimation of 7·28° for an angle of 60°; that is to say, the angle of 60° is judged to subtend 67·28°—to an underestimation of 4·74° for an angle of 140°; thus, the angle of 140° is judged to subtend 135·26°. All first-quadrant, acute angles are *overestimated* and this distortion proceeds through the fourth quadrant to an angle of 110°, whereupon the pattern changes to one of underestimation. This trend is maintained to angles of 165°; subsequently reverting to a further small area of overestimation. All the distortions are significant, in the sense of differing statistically from actual physical size, with the exceptions of those referring to angles of 90°, 110°, 115°, 155°, 170°, 175° and 180°. It may be noted that of a total of almost five thousand responses made by subjects in this experiment, only three were not given to the nearest 5°. Hence, although angular judgements are undoubtedly most accurate at the extremities and in the centre of the range studied, judgemental rounding probably results in overestimation of accuracy in these regions.

Experiment 2

The results of Experiment 1 reveal that acute angles are overestimated in size throughout, while obtuse angles are underestimated over a wide range. Thus, they provide strong evidence in favour of the statements made by both Helmholtz and Wundt but they appear to be contrary to the findings of Jastrow. Most previous observations have indicated that, in general, acute angles are overestimated and obtuse angles underestimated. Innumerable statements of this kind have appeared in the literature for well over a hundred years. Nevertheless, *they cannot* *be true*. The statement "*All acute angles are overestimated*" must necessarily be in error. Consider a triangle. In Euclidean geometry this is defined as a three-sided, rectilinear figure, the sum of the angles included within which is 180°. The three intersecting lines of an equilateral triangle subtend 60°. Thus, they are all acute. Moreover, two 45° angles are contained within an isosceles, right-angled triangle and these too are acute. If the sizes of all acute angles were to be overestimated it should follow that triangular figures are distorted so that their sides assume a bowed appearance. Since triangles appear to preserve their characteristic shapes and their sides to remain straight, it must follow that *all* acute angles cannot be overestimated in size.

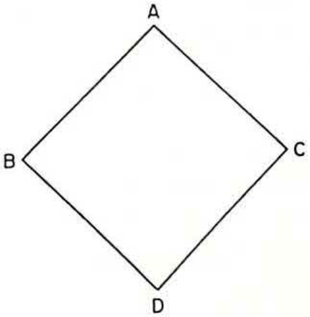


FIGURE 3. The Schumann version of the vertical-horizontal illusion.

Although the sides of triangular figures do not appear to bow certain features of them are susceptible to distortion. A triangle, of the same base and height, usually appears taller than it is wide. This is one example of the Fick (1851), or vertical-horizontal, illusion. Under the influence of this illusion the vertical dimension of any rectilinear, regular and symmetrical figure, identical in height and base, is judged to exceed that of its horizontal extent. While being disarmingly simple this particular illusion has proved extremely resistant to explanation. A further example of apparent-distortion of this same kind appears in Schumann's (1900) variant upon the vertical-horizontal illusion shown in Figure 3. The illusory effect is such that the distance AD appears greater than that of BC. Since this figure is actually a square rotated about the upright through an angle of 45°, however, its diagonals are identical in length. The diamond-shaped appearance implies that the four angles formed by adjacent sides are distorted. To maintain internal formal consistency and to preserve the rectilinear nature of Figure 3, the left- and right-pointing right-angles ABD and ACD must be *overestimated* in size, while the upward- and downward-pointing angles BAD and BDC must be *underestimated*. The geometrical equilibrium of this rectilinear figure can only be preserved if the angles included within it are distorted in different directions according to their orientations.

Method and procedure

From the materials prepared for use in Experiment 1, four further sets of photographic slides were made. One of these was arranged so as to display the angles of 5° through 180° facing to the left symmetrically about the horizontal axis (which was not included); the second set faced to the right; while the third and fourth sets faced upwards and downwards symmetrically about the vertical axis. This whole set of 144 slides was exposed as before to the same sample of 132 individual subjects, each of whom was seated in the specially constructed chair described above. The random order in which the stimuli were presented was changed for each participant and all subjects were required to make an absolute judgement, in degrees, of the size of each angle.

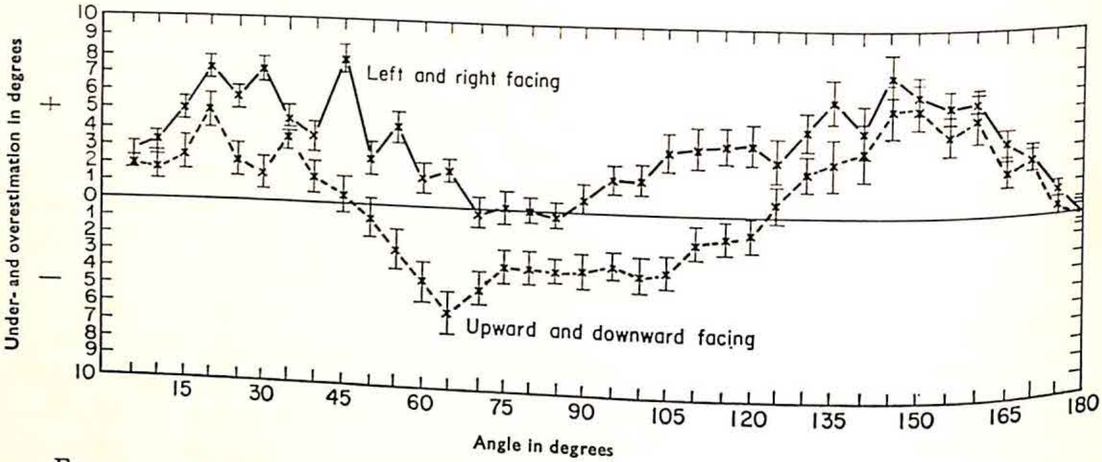


FIGURE 4. Absolute judgements of the sizes of angles in different orientations.

Results

The results of this experiment are shown in graphical form in Figure 4, using the same convention as before. The apparent-sizes of otherwise similar left- and right-pointing angles were found not to differ significantly. Therefore, the data referring to them were merged and they are plotted in solid line. Similarly, upward- and downward-pointing angles did not differ and the judgements referring to them were pooled and plotted in broken line. These graphs indicate clearly that the usual statements describing apparent-distortion of acute angles and that the data obtained in Experiment 1 are of limited generality. With the exception of the statement made by Jastrow, which will be considered in detail in the next experiment, obtuse angles are usually described as underestimated in size and acutes as overestimated. When angular forms are presented in different orientations, however, this pattern undergoes a marked change. First, without exception, the sizes of all sideways-pointing angles are judged in excess of those which point upwards and downwards. This difference is statistically significant throughout the majority of the range included. In general, sideways-pointing angles are overestimated but a region of highly accurate judgements appears between 70° through 90°. Secondly, upward- and downward-pointing angles are underestimated over a wide range. Here, however, the distribution of judgements is less straightforward. Overestimation characterises judgements of angles below 45° and the angle of 45° itself is judged accurately. Subsequently, a broad area of underestimation appears up to and including 120°, whereupon a further range of overestimation becomes apparent.

The outcome of this particular part of the study is relevant to the mechanism underlying judgements of angular sizes. An angle is formed by two lines which differ in inclination. Perhaps the most obvious basis for making angular judgements would be for the apparent-inclinations of the two component lines to be determined first, whereupon appropriate computation should reveal the size of the angle included within them. The data show that the angle of 90°, for example, is slightly overestimated when facing to either left or right but that it is underestimated when pointing upwards or downwards. Statistically, the difference between these two judgements is significant ($t = 3{\cdot}48 : P < 0{\cdot}01$). The lines which compose the four 90° angles being considered, however, have identical inclinations; that is to say, the left- and right-pointing angles are both formed by two lines inclined at angles of 45° to the vertical. So are the angles pointing upwards and downwards. Since the sizes of these angles are judged differently depending upon their orientations, however, it must follow that the mechanism responsible for apparent angular distortion cannot operate in this simple way.

The prediction concerned with judgements of angles seen in different orientations, developed from the distortion seen in the Schumann figure, may be understood as receiving partial support by these results. The outcome is not entirely as expected, however, since the two graph-lines of Figure 4 should be exact mirror-images of each other; this being the only way in which the formal consistency of the diamond-shaped figure can be maintained. Presumably, therefore, isolation of the angular components of the Schumann figure from the whole form is instrumental in modifying its configurational properties in some hitherto undetermined way.

Experiment 3

The outcome of Experiment 1 seems to confirm that acute angles are overestimated and obtuse angles underestimated. The only significant exception to this finding was made by Jastrow. The volume of study undertaken by Jastrow in many fields of perception indicates that he was an extremely competent experimental psychologist of his time. Accordingly, it would be unreasonable to dismiss his findings without first seeking possible reasons for his contrary view. Describing the procedure adopted in his experiment Jastrow wrote "*. . . The subject viewed the angle as long as was needed to fix it in his mind (from* 5 *to* 15 *sec.) and immediately thereupon, from his memory of the angle, drew another as nearly as possible equal to the first . . .*" (1892, p. 214). Consider the usual statement describing angular distortion. An acute angle is *seen, and thus judged,* to be larger than it actually is. Hence, when a subject is asked to make an absolute judgement indicating the apparent-size of a given acute angle, he will *report it to be in excess of its physical size.* Suppose, however, that he is now asked to *construct an angle of specified size* by drawing two lines in appropriate juxtaposition. *Since the angle is judged to be larger than it is, when it is reproduced by construction, appropriate lines will be drawn subtending an angular distance smaller than the angle specified.* The purpose of the third experiment was to submit this prediction to experimental consideration.

Apparatus and procedure

A circular piece of dense white plastic material, 300 cm. in diameter, was arranged in such a way as to provide a near-infinite visual field for subjects seated in the specially constructed chair at a distance of 30 cm. A black circular spot, 5 mm. in diameter, was placed in the centre of the field and a black line, 45 cm. long and 2 mm. wide was painted so as to extend from the central point horizontally to the right. This was intended to provide a datum reference of 0°. The subjects were equipped with a transparent, ungraduated straight-edge and a wax pencil with which they were required to construct angles ranging from 20° to 340°, in increments of 20°. They constructed the angles of 45°, 90° and 270° also. All the angles were reproduced in a randomly determined order and this was changed for every participant subject, who constructed each angle once only.

Results

The results of this experiment are shown in Figure 5. They refer to the judgements made by 145 subjects none of whom had taken part in either of the two previous studies. It was suggested above that if acute angles are seen as larger than their physical sizes then specified angles should be constructed commensurately smaller. These results reveal unquestionably that this is the case and they are tolerably in agreement with the results of Experiment 1. According to the data shown in Figure 2, however, the area of significant underestimation of obtuse angles is confined to the range 115° to 165°. A similar range of overestimation is evident in Figure 5 but it fails to parallel the previous data exactly. No comparison data are available for angles beyond 180° but the symmetry of the further two cycles of the graph seems intuitively to be in keeping with the first two.

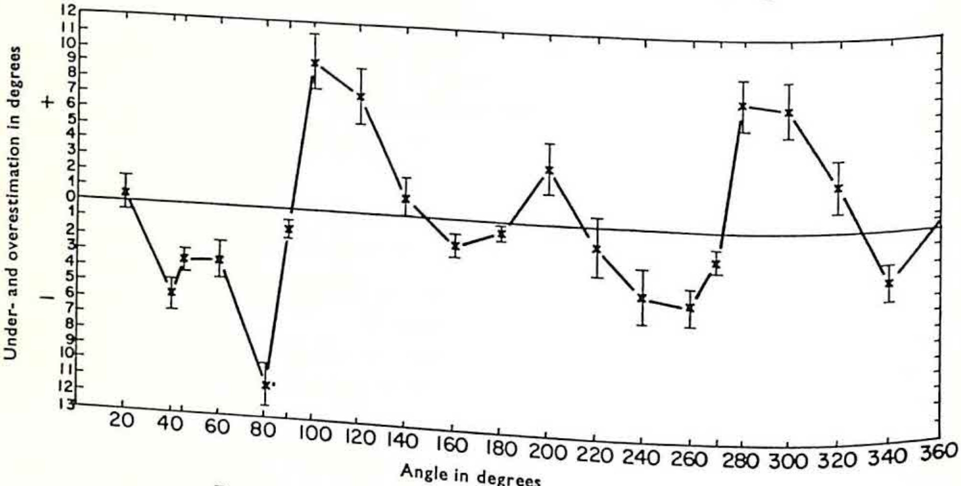


FIGURE 5. Constructing angles of specified sizes.

The data illustrated in Figure 5 may be compared with the results obtained by Jastrow, shown in Figure 1. He observed an area of underestimation for acute angles which might be regarded as more-or-less in keeping with the present findings but important reservations to such an interpretation will be developed in discussion. Beyond 90° Jastrow found an area of overestimation for absolute judgements and this reverted to underestimation for angles in excess of 135°. Again, this is confirmed tolerably well by these results where the reversal occurs at 140°. Accordingly, it may be concluded that a broadly consistent relationship has been demonstrated between absolute judgements of angular size and those involved when given angles are constructed.

Discussion

Apparent-distortion of angular size has been discussed widely in studies of visual perception. It has been referred to principally for its relevance to visual illusions and, as pointed out previously, the problem of why angles themselves appear to undergo distortion has been largely ignored. Despite the volume of study, much lack of agreement remains over the apparent-sizes of angles. Hering (1861, p. 65) reported that the threshold for over- and underestimation occurs at an angle of 60°, whereas most writers consider the changeover to take place at 90°. Helmholtz frequently subscribed to the view that acute angles are overestimated in general, but he sometimes restricted this to patterns in which they are juxtaposed with obtuse angles. Wundt was aware of the phenomenon of apparent angular distortion but considered it of secondary importance to the facility with which movements of the eyes can be initiated in different directions. He would almost certainly have interpreted the outcome of Experiment 2 in accordance with this view. Judd (1899) also claimed that angular distortion was a secondary effect, in this case to judgements of the relative lengths of associated lines, as in the Ponzo (1912) illusion. Ladd and Woodworth attempted to dismiss the whole question of the relevance of apparent angular distortion altogether in claiming that it is "*. . . doubtful whether the overestimation of acute angles is, in reality, a universal principle which can be fully invoked for the explanation of these illusions . . .*" (1911, p. 445). Pratt found that "*. . . both acute angles and obtuse angles* (45° *and* 135° *were taken to be typical*) *were as apt to be overestimated as they were to be underestimated . . .*" (1926, p. 140) and, hitherto, Samanta (1928) would probably have been regarded as the sole supporter of Jastrow's view since he reported that acute angles are *underestimated* in size.

Three factors suggest, however, that distortions of angular size are not of such limited generality as to discount them entirely. These are (i) the pattern of distortion evident throughout the range of first and fourth quadrant angles; (ii) the systematic tendencies for angles to be differentially distorted in accordance with their orientations; and (iii) the relationship between absolute judgements of angular subtension and the sizes in which specified angles are constructed. Moreover, the results of the present experiments indicate that all previous conflicting findings may be understood in terms of the appropriate situation described here.

At first sight it might appear that Jastrow's findings could be accommodated by the present results also. The procedures adopted in Experiment 3 were intended to establish the relationship between absolute judgements of the sizes of angles and constructing angles of given sizes. The observations of Helmholtz and Wundt were confirmed in terms of absolute judgements, while those of Jastrow appeared to be in accordance with the sizes of the angles constructed. It should be emphasized, however, that although the procedures used in this experiment were very similar to those employed by Jastrow, the two situations are not strictly comparable. From his description of the experiment which he devised, it seems clear that he required his subjects to construct given angles subsequent to having made judgements of their sizes; these presumably having been expressed, implicitly or explicitely, in absolute terms. Thus, having judged a given acute angle to be in excess of its actual size, the subjects proceeded to reproduce it. But it is difficult to see how Jastrow could have made any prediction as to distortions expected to arise in this situation. Suppose, for example, that when shown a 40° angle a subject judges it to be one of 45°. When attempting to reproduce it he will construct two lines *which he believes to represent an angle of* 45°. What, therefore, will be the size of the ensuing angle? The outcome of Experiment 3 indicated that a *given* angle is constructed smaller in size than the one required. But in Jastrow's study the size of the angle was judged rather than specified. Without having first evaluated the distortions arising in these two quite different situations, the only prediction open to Jastrow was that *no distortion should have become apparent* in the experiment which he performed.

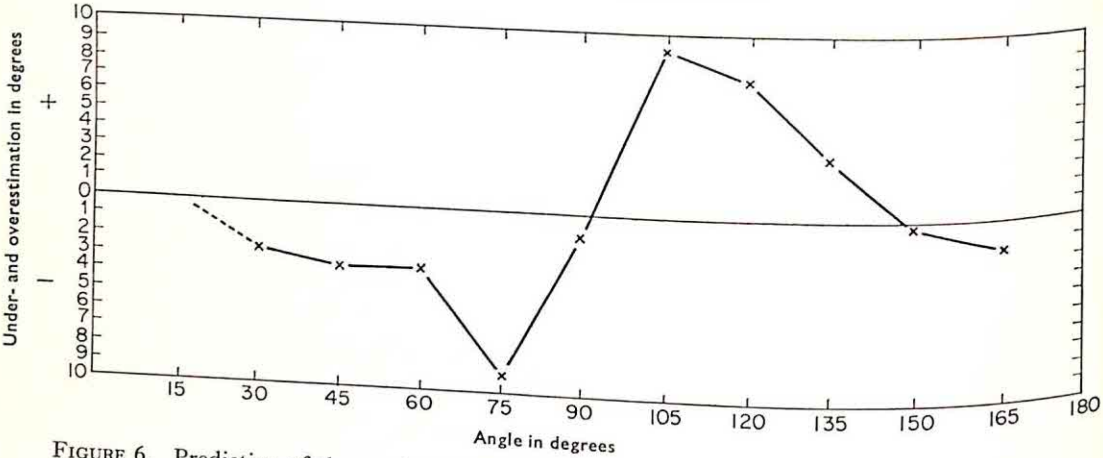


FIGURE 6. Prediction of the results of Jastrow's experiment from the data obtained in Experiments 2 and 3.

It is possible that this problem could be ameliorated, however, by further comparison of the results of the present study and those obtained by Jastrow. In his experiment, having made absolute judgements of the sizes of angles, the subjects were required to reproduce them. Errors in absolute judgements are shown in Figure 2. This provides the best estimates available of the sizes of the angles which Jastrow's subjects actually attempted to construct. Approximations to the errors involved in reproducing *these particular angles rather than the angles specified* may be obtained by interpolation from Figure 5. Had Jastrow carried out these two experiments prior to his major study he would have been in a position to predict its outcome. The appropriate values have been plotted in Figure 6. Comparison of these with Figure 1 indicates a somewhat more reasonable fit than was obtained from comparison of Jastrow's results with the outcome of Experiment 3.

The mistaken logic of Jastrow's experiments arises frequently in scientific discussion. It is exemplified by the reasons suggested by art historians and others for the distortions which characterize the works of certain artists. Schepler, for example, observed "*You can tell an artist's visual defects by looking at his works of art. The Renaissance painters Holbein, Cranach and El Greco each had a form of astigmatism. Holbein saw everything flattened from side to side, Cranach saw them* *flattened up and down, and El Greco saw them oblique. His figures gave the impression of leaning towards the right. When viewing the paintings of these artists through proper corrective lenses, all the figures become normally proportioned.*" (1968, p. 8). Thus, it is commonly supposed that these artists suffered from visual defects of some kind. Had they intended their techniques to be representational, however, they would have attempted to portray their subject faithfully; that is to say the actual and reproduced forms would have *seemed to them* to be more-or-less in correspondence. Since any systematic distortions of the subjects seen by the artists would have been countermanded when they attempted to reproduce them, it is clear that the characteristic styles of their paintings fail to be explained by any defects of their vision.

Perhaps the most important argument developed in this paper rests upon demonstrating a consistent relationship between absolute judgements of angular subtension and the judgements involved when specified angles are constructed. As far as I am aware, a relationship of this kind has not previously been proposed. Wundt attempted to explain a number of examples of apparent spatial distortion by the facility with which the eyes can be moved in different directions. He gave particular attention to the vertical-horizontal illusion which lends itself to an explanation expressed in these terms. Wundt also attempted to explain why angles appear distorted when they are examined tactually. He adopted a principle similar to that which he invoked to account for illusions perceived visually observing "*. . . The tactual analogues of the illusion in visual angles is to be found in the tendency to overestimate small articular movements and to underestimate large ones. This comes under the general principle that a relatively greater expenditure of energy is required for a short movement than for a more extensive one . . .*" (1902, p. 137). The experiments of Robertson (1902), Révész (1934), Over (1966) and Fisher (1966) indicate that illusory distortions are not confined to vision. They are experienced when spatial information is impressed upon the skin and when displays are examined by the fingers and thumb in the absence of vision. When Pritchard (1958) stabilized the images of the Zöllner (1860) and Luckiesh (1922) illusions on the retina he reported that they appeared as clearly as in normal vision. It follows that visual illusions cannot be explained by eye movements and it seems unlikely that tactual distortions will be explained by the facility of moving the hands.

References

FICK, A. (1851). *De errore quodam optico asymmetria bulbi effecto.* Marburg: J. A. Koch.
FISHER, G. H. (1966). A tactile Poggendorff illusion. *Nature, Lond.* **212**, 105–6.
GUILFORD, J. P. and DALLENBACH, K. M. (1928). A study of the autokinetic sensation. *Am. J. Psychol.* **40**, 83–91.
VON HELMHOLTZ, H. (1962). *Handbuch der physiologischen Optik*, III, trans. by SOUTHALL, J. P. C. (ed.). *Helmholtz's Treatise on Physiological Optics.* New York: Dover Publications.
HERING, E. (1861). *Beiträge zur Physiologie*, I. Leipzig: W. Engelmann.
JASTROW, J. (1892). On the judgement of angles and positions of lines. *Am. J. Psychol.* **5**, 214–48.
JUDD, C. H. (1889). A study of geometrical illusions. *Psychol. Rev.* **6**, 241–61.

LADD, G. T. and WOODWORTH, R. S. (1911). *Elements of Physiological Psychology*. New York: Scribner's.

LUCKIESH, M. (1922). *Visual Illusions: Their Causes, Characteristics and Applications*. New York: Van Nostrand.

OVER, R. (1966). A comparison of haptic and visual judgements of some illusions. *Am. J. Psychol.* **79**, 590–5.

PONZO, M. (1912). Rapports de contraste angulaire et l'appréciation de grandeur des astres à l'horizon. *Arch. Ital. d. Biol.* **58**, 327–9.

PRATT, M. B. (1926). The visual estimation of angles. *J. exp. Psychol.* **9**, 132–40.

PRITCHARD, R. M. (1958). Visual illusions viewed as stabilized retinal images. *Q. J. exp. Psychol.* **10**, 77–81.

RÉVÉSZ, G. (1934). System der optischen und haptischen Raumtäuschungen. *Z. Psychol.* **13**, 296–375.

ROBERTSON, A. (1902). Studies from the psychological laboratory of the University of California: VI 'Geometrico-optical' illusions in touch. *Psychol. Rev.* **9**, 549–69.

SAMANTA, M. N. (1928). Visual estimation of angles. *Indian J. Psychol.* **3**, 185–8.

SCHEPLER, H. (1968). The "eyes" have it. *Spie Glass* (Society for Photo-optical Instrumentation Engineers), **4**, 8.

SCHUMANN, F. (1900). Beiträge zur Analyse der Gesichtswahrnehmungen. *Z. Psychol.* **23**, 1–32.

WUNDT, W. (1902). *Beiträge zur Theorie der Sinneswahrnehmung*, trans. by JUDD, C. H. *Outlines of Psychology*. Leipzig: W. Engelmann.

ZÖLLNER, F. (1860). Über eine neue Art von Pseudoskopie und ihre Beziehungen zu den von Plateau und Oppel beschriebenen Bewegungsphänomenen. *Ann. Phys. Chem.* **186**, 500–20.

Received 9 March 1969

Q. Jl exp. Psychol. (1969) **21**, 367–380

SOME EFFECTS OF FRUSTRATION IN A FREE RESPONDING OPERANT SITUATION

R. N. WILTON,† K. T. STRONGMAN

University of Exeter

A. NERENBERG

University of Texas

In several experiments, rats and pigeons were reinforced on a chained schedule which varied in its second component but which always had a variable interval schedule as its first component. Occasionally, the reward which normally followed responding in the second component was omitted. On these frustrative non-reward occasions responding immediately subsequent to non-reward was sometimes elevated and sometimes depressed. The relevance of the elevation and depression effects to behavioural measures of frustration in the double runway was discussed.

Introduction

A frustrative event has been defined (Amsel, 1958) as the encountering of non-reward in a situation in which reward was previously found. In a free responding operant situation an omission of reward following consistent reinforcement constitutes one case of frustrative non-reward which may lead to an immediate increase in response vigour. Extinction, caused by the continued omission of reward is often characterized by an initial increase in response rate (Skinner, 1938; Keller and Schoenfeld, 1950). In the present experiments, a technique was designed to generate this effect, and allow other effects of frustrative non-reward to be observed, without the subsequent decline in rate due to continued non-reinforcement. The second component of a two-component chain schedule was occasionally terminated without reward. Effects on immediately subsequent behaviour were then observed, while overall responding was maintained by the reward obtained on the remaining occasions, when the second component was allowed to run its full course.

Experiment 1

In Experiment 1, rats were trained to respond on a modified variable interval schedule. A tone was contingent upon variable interval responding, and on 50 per cent of its presented occasions this tone provided an occasion for reward. On the other occasions reward was not available; the rat experienced frustrative non-reward. The reward consisted of sucrose solution, and the concentration of this solution was varied during the experiment to determine whether post-frustration effects are related to the magnitude of the reinforcer which is omitted in the frustrating situation.

---

† Present address: University of Texas, Houston, Texas, U.S.A.

Method

Subjects

Four experimentally naive male hooded rats were housed in individual cages, and were approximately 4 months old at the outset of experimentation. Water was always available in each home cage, and 12 g. of solid food was given to each rat at the end of its daily experimental session.

Apparatus

One sound insulated operant conditioning chamber contained a lever and a reinforcement tray. The lever was mounted 2½ in. above the floor level. The 1½ in. wide × 2 × 2 in. reinforcement tray was recessed into the wall of the chamber, 2¼ in. to the left of the bar. A small fountain of sucrose solution could be licked by the rat, just below the level of a ¼ in. diameter hole in the floor of the tray. The fountain was produced by a pump, which could be operated for 1½ sec. by a break in a photobeam which crossed the reinforcement tray. A break in this photobeam constituted a tray response. The mechanical latency of the pump was long enough to allow a rat access to the sucrose solution for the full 1½ sec. exposure time. Sessions were programmed through relays, timers and print-out counters contained in an adjoining room. The reinforcement consisted of sucrose solution, made up by weight of sucrose to volume of water.

Preliminary procedure

The rats were first adapted to the chamber, and then trained to approach the reinforcement tray for a 4 per cent sucrose solution in the presence of a 0·75 sec., 50 Hz tone. The lever was then inserted into the wall of the chamber, and the rats were gradually shifted from consistent reinforcement to a 1-min. variable interval schedule (Fleshler and Hoffman, 1962). Also, the tone, contingent upon the lever response according to the reinforcement schedule, was now followed by an available reward on only a random 50 per cent of its presentations. Hence, a tray response following the tone operated the pump on reinforced occasions, and failed to operate it on frustrated occasions. The complete sequence may be viewed as a two component chain schedule (VI 1 FR 1 PR 50 per cent), in which a lever response (first component) occasionally produces a tone which is the discriminative stimulus for a tray response (second component), which is reinforced on a 50 per cent partial reinforcement schedule.

The termination of the first lever response, following a frustrated (non-reinforced) tray response, started a timer which allowed the number of responses emitted in the following 10 sec. (called a post-frustration period) to be recorded. A 10-sec. period of responding was also sampled every 2 min. throughout the session (a baseline period), again through a timer started by the termination of a lever response. If an occasion to produce the tone (VI schedule) became available within a response recording period, then the responses so far accumulated, and the responses subsequently emitted in that period, were erased. A total of 15 sessions was given, by which time performance had stabilized. Each session was terminated after the presentation of 50 reinforcements (about 100 min.). The rats were then provided with their daily ration of food.

Procedure

The procedure was split into two stages. In a shift up stage the concentration of sucrose solution was raised from 4 up to 32 per cent. In a shift down stage it was lowered from 32 back to 4 per cent.

Shift up

Forty-eight sessions were given, in the course of which the concentration of sucrose solution was raised from 4–8–16–32 per cent. Sucrose at each concentration served as a reinforcer for 12 sessions.

Shift down

Fifty-six sessions were given, in the course of which the concentration of sucrose solution was lowered from 32–16–8–4 per cent. The 16 per cent solution was given for 32 sessions, when the periods recording post-frustration and baseline responses were altered in duration from day to day in the order 3, 5, 7, 10 and 15 sec. By doing this it was hoped to assess the duration of frustration produced energizing effects as a function of time elapsed since the onset of a post-frustration period. The 8 and 4 per cent solutions each served as a reinforcer for 12 sessions, the response sampling periods then being set at the standard 10 sec.

Results

Responding following a shift in a sucrose concentration often stabilized only after several sessions. For this reason, data from the first seven sessions on a new concentration were discarded. All subsequent sessions are taken into account in the following analyses.

Figure 1 shows baseline rate and post-frustration rate at each concentration of sucrose solution. Each point represents the mean rate over five sessions. Inspection of Figure 1 shows that the post-frustration rate is consistently above the baseline rate. Also, the difference between the post-frustration and baseline rates tends to increase with increases in sucrose concentration. A two-tail *t*-test indicates that the difference between the post-frustration rate and the baseline rate at the 32 per cent concentration is greater than the difference at the 4 per cent shift up concentration ($P < 0{\cdot}04$), and greater than the difference at the 4 per cent shift down concentration ($P < 0{\cdot}04$). Changes in the difference between post-frustration rate and baseline rate do not appear to be particular to either the shift up or shift down stages, though all changes could depend on experience of the shift up stage.

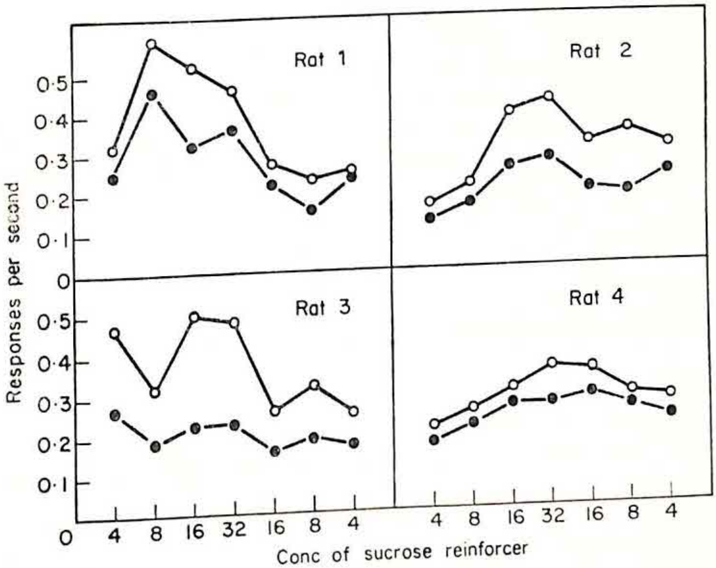


FIGURE 1. Mean rate of responding for each of the four rats, taken over five sessions at each concentration in the shift up and shift down stages of Experiment 1. (○) Mean post-frustration rate, (●) mean baseline rate.

There is evidence of a further post-frustration effect, the occurrence of which seems peculiar to the 32 per cent concentration. Figure 2 shows a frequency distribution of post-frustration periods, classified according to the number of contained responses. These data are taken only from the shift up stage of the experiment. Similar data were not available for the shift down phase. At each concentration, the periods containing the maximum number of responses generally contain more responses than the periods containing the maximum number of responses at lower concentrations. Thus, with increases in concentration, there is a trend towards an increase in the maximum number of responses emitted in a post-frustration period. There is also a trend towards a lower frequency of periods containing a very few (zero or one) responses. But this trend is reversed at the 32 per cent concentration. There is a marked increase in the frequency of periods containing either a single response, or no responses, at the 32 per cent

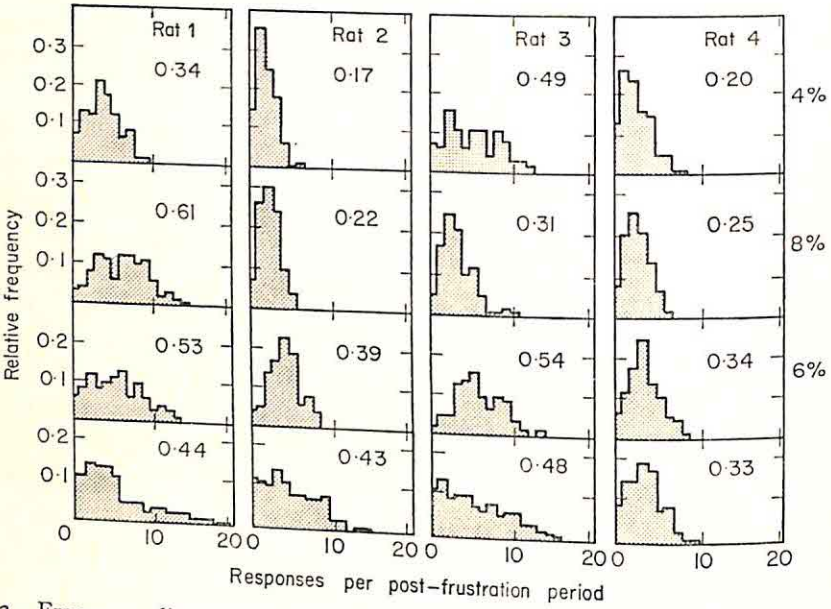


FIGURE 2. Frequency distribution of post-frustration periods classified according to the number of contained responses. The data are obtained from five sessions at each concentration in the shift up stage of Experiment 1. The mean post-frustration rate over the five sessions is indicated.

TABLE I

Percentage frequency of post-frustration periods containing zero responses at the shift up concentration in Experiment 1

	4%	8%	16%	32%
R1	7	3	6	11
R2	14	7	4	12
R3	7	9	2	11
R4	13	8	7	9
M	10	7	5	11

Each score is the mean of scores from five successive sessions.

as compared with the 16 per cent. concentration. The total effect is that the obtained frequencies up to 32 per cent have a fairly symmetrical distribution, but at 32 per cent they show positive skew. Table I shows the relative frequencies of periods containing zero responses at each concentration. Each cell at the 16 per cent concentration contains a lower score than a corresponding cell at both the 4 and 32 per cent concentrations (two-tail *t*-test, $0{\cdot}05 > P < 0{\cdot}10$, $P < 0{\cdot}04$).

A histogram, like that in Figure 2, which shows an increase in the frequency of periods containing very few responses, could be obtained if there was rapid responding at the beginning of a session, and slow responding towards its end. In order to test this hypothesis, the frequency of periods containing zero responses was determined for each quarter of a session. Table II shows that in the first two quarters of a session all the scores at 32 per cent are above the corresponding scores at 16 per cent. To account for these data the satiation hypothesis must suppose an extremely rapid onset of satiation at the 32 per cent concentration, while also accounting for a high overall rate (Fig. 1). Also, the frequency of periods containing zero responses tends to be high in the first quarter of a session, lower in the second and third quarters, and high in the fourth quarter of a session. The satiation hypothesis would have difficulty in explaining the drop in frequency of periods containing zero responses from the first to the second and third quarters of the session; but it would have no difficulty in accounting for the rise in frequency from the third to the fourth quarter. If frustration is to be held at all responsible for the effect, it would most likely account for the effect in the first quarter of the session. And so, perhaps there is some indication that some effects of frustration are less likely to occur as the session progresses.

TABLE II

Percentage frequency of post-frustration periods containing zero responses in each quarter of a session at the shift up 16 *and* 32 *per cent in Experiment* 1

	1st		2nd		3rd		4th	
	16%	32%	16%	32%	16%	32%	16%	32%
R1	2·5	8·0	2·5	8·0	10·0	8·0	8·0	16·0
R2	5·0	22·0	0·0	4·0	7·5	6·0	0·0	15·5
R3	2·5	20·0	0·0	4·0	6·0	4·0	0·0	16·0
R4	12·5	15·0	2·5	5·0	10·0	7·5	0·0	16·5
M	5·6	16·3	1·3	5·3	8·4	6·4	2·0	16·0

Each score is the mean of scores from five successive sessions.

Figure 3 contains the data from the shift down stage of the experiment, when the reinforcer was a 16 per cent sucrose solution, and the periods recording post-frustration and baseline responses were systematically varied from 3 to 15 sec. The data are ordered in the form of cumulative records. The open circles represent baseline responses. The baseline responses tend to accumulate regularly as the time since the onset of a baseline period increases, indicating a fairly consistent baseline response rate. The filled circles represent the difference between post-frustration and baseline responses. Responses contributing to the difference cease to accumulate after about 10 sec. for rats 1 and 4, and some time after 15 sec. for rats 2 and 3.

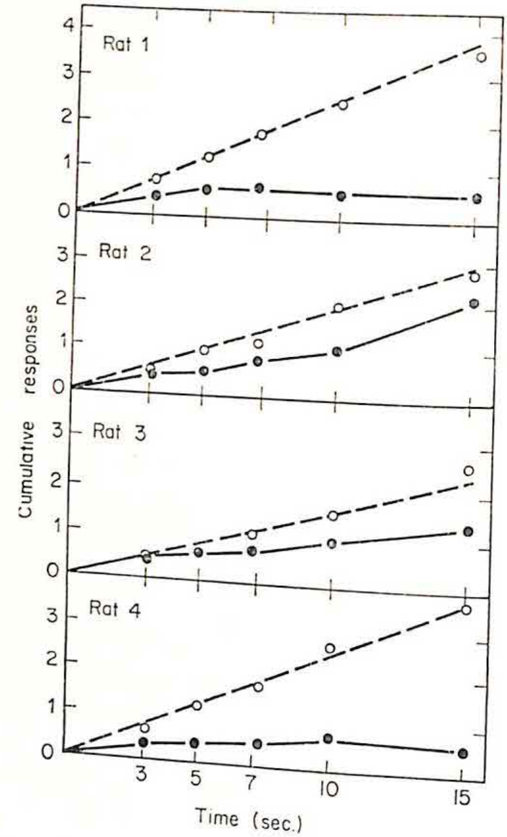


FIGURE 3. Mean number of responses emitted at the 16 per cent shift down concentration in Experiment 1. Responses are plotted as a function of time elapsed since the onset of a post-frustration or baseline period. Each point represents data from five sessions. (●) Difference between post-frustration and baseline responses, and are obtained by subtracting baseline responses (○) from total post-frustration responses.

Discussion

In the present experiment, post-frustration responding is found to be elevated above post-non-frustration responding. Also, there is some tendency for this elevation to increase with increases in the magnitude of the reinforcer omitted in the frustrating situation. If the chosen post-non-frustration rate is accepted as a "correct" baseline measure, then it seems that the post-frustration elevation can be identified with the frustration effect (FE) found in the double runway (Wagner, 1959), or the increase in response rate, or vigour, frequently found in the free responding operant situation during the early stages of extinction, when extinction follows a history of consistent reinforcement (Skinner, 1938; Keller and Schoenfeld, 1950; Thompson, Heistad and Palermo, 1963). Although the increase in response rate, occurring in the free responding situation, is generally assumed to be due to emotional or frustrative effects of non-reward, this has yet to be conclusively demonstrated. More usually, it could be due to an absence of time spent consuming food, an absence of satiation effects, or an absence of some other effect of reinforcement. In the present study, there may be grounds for questioning the appropriateness of the chosen baseline measure, but it should be noted that a post-reinforcement rate measure would not be a good alternative. The conditions preceding post-reinforcement responding differ from those preceding post-frustration responding not only because they include the absence of frustration, but also because they include the presence of reward. Reinforcement could have its own immediate effects, either excitatory (Maltzman, 1950), or inhibitory (Hamm, 1967). To compare post-frustration responding with post-reinforcement responding would be to confound the effects of frustration with the effects of reinforcement.

A comparison of sessional post-frustration rates at different sucrose concentrations indicates a general trend for rates to increase with increases in concentration (Fig. 1), and the results at the 32 per cent concentration are in accordance with this trend. However, a more molecular analysis (Fig. 2) shows a new effect at the 32 per cent concentration, post-frustration responding is occasionally attenuated. The frequency of post-frustration periods containing very few responses is greatly increased at the 32 per cent as compared with the 16 per cent concentration. The occurrence of this effect in the first half of the session seems to preclude an interpretation based on satiation.

Skinner (1950) reports a similar effect. A pigeon was trained to peck at a disc on the presentation of a specific stimulus, and its response latency was recorded. When reinforcement was discontinued, a scattering of latencies occurred, some being very short, others very long. Skinner attributes the occurrence of the very long latencies to response interference by emotional behaviour (and specifically not to response extinction). Figure 4 illustrates this effect. In the present terminology the upper portion of the figure indicates baseline latencies (actually post-reinforcement), and the lower portion indicates post-frustration latencies. In this within reinforcement condition an increase in the frequency of long

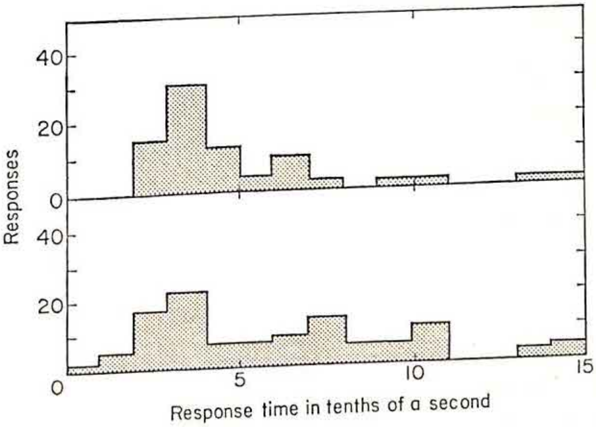


FIGURE 4. Pigeon's response latencies. Frequencies of particular latencies are shown under conditions of consistent reinforcement and no reinforcement (adapted from Skinner, 1950).

latencies appears in the post-frustration period. The figure can be most usefully compared with Figure 2, where post-frustration rates are compared between reinforcement conditions. The appearance of periods containing very few responses, the analogue of the long latencies, here characterizes the difference between responding in the 32 per cent and the lower concentration post-frustration periods. The similarity of the effect in the two studies supports the contention that the present post-frustration effects have common causal antecedents with those which precede normal extinction.

The technique used in the present experiment does seem to produce post-frustration effects which can be constantly regenerated. Given this, one modification might increase its flexibility. The substitution of a more standard form of response, such as a second lever, for the tray response, would enable frustrative non-reward to be programmed easily when the second component of the multiple schedule consists of a reinforcement contingency other than FR 1. In the two following experiments, pigeons are used as experimental organisms, and key peck responses replace the lever and tray response.

Experiment 2

In this experiment three pigeons were reinforced on a chain schedule, similar to that used in Experiment 1. A peck at one key (first component) was equivalent to the lever press of Experiment 1, and a peck at a second key (second component) was equivalent to the tray response of Experiment 1. The occasion for responding to the second key was indicated by its illumination, which was contingent upon a response to the first key in the same way that the tone used in Experiment 1 was contingent upon a lever response.

Method

Subjects

Three experimentally naive adult homing pigeons were maintained at about 80 per cent of their free feeding weights.

Apparatus

One sound insulated operant conditioning chamber contained two translucent response keys of ¾ in. diameter. One key could be illuminated from behind by a white light, the other by a red light. A grain magazine was centred below the white key. Reinforcement consisted of 4½ sec. access to a grain mixture. Sessions were programmed through relays and timers contained in an adjoining room.

Preliminary procedure

The birds were first adapted to the chamber, magazine trained, and trained to peck at one of the keys when it was occasionally illuminated by a white light. A first response to the illuminated white key darkened it, and operated the grain magazine. A chain schedule was then introduced whereby a response to a second key (illuminated red) illuminated the first key on a probability governed by a 1 min. variable interval (VI 1) schedule. This was the same schedule as the VI schedule used in Experiment 1 (Hoffman and Fleshler, 1962). When the first (white) key was illuminated, the second (normally red) key darkened, and was illuminated again only when a response to the white key operated the grain magazine. The sequence constituted a two component chain schedule (VI 1 FR 1), in which a response to a red key (first component) occasionally illuminated another key (second component), a response to which operated the grain magazine. Daily sessions of 50 min. each followed an initial period of adaptation to the schedule. Response rates were stable after 20 sessions. The experimental procedure then began.

Procedure

The procedure was split into two stages. In each stage the daily session length was 100 min.

Stage 1. The schedule was now altered so that a response to the illuminated white key operated the grain magazine on only a 50 per cent random basis. Thus, a response to the white key now either operated the grain magazine (reinforcement), darkened the white key, and returned the bird to the VI schedule, or failed to operate the grain magazine (frustration), darkened the white key, and returned the bird to the VI schedule. The termination of the first response to the red key following frustration started a timer which allowed the number of responses emitted to the red key in the next 10 sec. (post-frustration

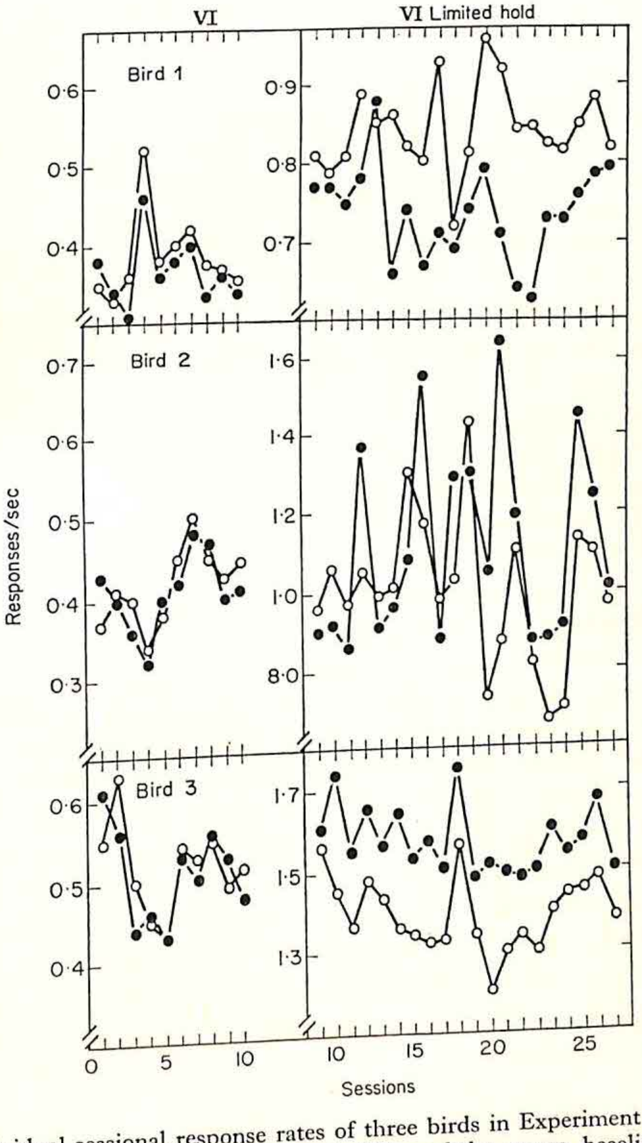


FIGURE 5. Individual sessional response rates of three birds in Experiment 2. Response rates are taken from the 10-sec. post-frustration periods (●), and the 10-sec. baseline period (○).

period) to be recorded. Ten second periods (baseline periods) of responding were also sampled throughout the session, again through a timer, started by the termination of the first response to the red key which followed a 2-min. interval since the last baseline period. Both timers failed to operate if a "reinforcement" (occasion to illuminate the white key) was already set up on the red key when the first (start timer) response to the red key was emitted. A total of 10 sessions was given.

Stage 2. Conditions were the same as in stage 1, except that a limited hold of 0·9 sec. was imposed on the VI schedule of the first component (red key). This meant that each "reinforcement" (occasion to enter the second component) was available for only 9 sec., and was missed if the bird failed to respond to the red key within that time. In the first five sessions an initial limited hold of 2 sec. was gradually reduced to the final 0·9 sec. A total of 28 sessions was given.

Results and discussion

Figure 5 shows the response rates of each bird when the post-frustration and baseline rates are taken from individual sessions. When exposed to the ordinary variable interval (stage 1), none of the birds showed an elevation of post-frustration rate above baseline rate. There was, therefore, no evidence of energizing effect following frustration, and the results from the similar procedure of Experiment 1 which used rats as subjects, were not replicated. However, visual inspection of the birds showed that they tended to "pace" the experimental chamber following a non-reward response to the white key (see Azrin *et al.*, and Terrace, 1966, for similar effects).

The results may be compared with those obtained by Staddon and Innes (1966). An apparent frustration produced energizing effect here occurred on a fixed interval schedule. Post-frustration performance was compared with post-reinforcement performance. At present it would be difficult to say whether this was a genuine energizing effect produced in more favourable conditions, or a depression due to temporary satiation affecting the post-reinforcement rates.

A limited hold was imposed on the VI schedule in an attempt to reduce the incidence of behaviour other than that of responding to the red key. A limited hold typically leads to the development of rapid responding on a variable interval response key (Ferster and Skinner, 1957). In the present experiment it was hoped that the limited hold contingency would increase orientation to the red key, with the result that any energizing effects of frustration would be channelled into responding on the red key. Figure 5 shows the response rates of the three birds after eight sessions on the limited hold schedule. Birds 2 and 3 show some consistent elevation of post-frustration rate above the now increased baseline rate. Bird 1 shows an elevation of baseline rate above post-frustration rate, an exaggeration of the relationship shown previously on the ordinary VI schedule. Each bird showed a very irregular response rate within an experimental session.

It must be concluded that the energizing properties of frustration were not at all apparent in stage 1 of the experiment, and were not reliably or convincingly demonstrated in stage 2, although the occurrence of pacing did indicate emotional effects.

Experiment 3

In Experiment 3, the FR 1 component of Experiment 2 was replaced by a fixed interval (FI) component. This fixed interval could be terminated at various times before it had run its full course. This operation is similar to the frustrative non-reward operation of Experiments 1 and 2 in that a stimulus which indicates that responding may be rewarded (white key in Experiment 2, early part of FI in Experiment 3) is replaced by a stimulus which indicates responding will be unrewarded (absence of magazine sound in Experiment 2, absence of later part of FI in Experiment 3).

As before, post-frustration and baseline response rates were measured. Also measured was the time elapsed between the termination of the FI and the resumption of responding in the VI component.

Method

Procedure

The procedure was basically the same as that of Experiment 2, but was changed in the following ways. The colour of the VI key, previously red, was now blue. The FR 1 component was changed to FI 36 sec., and the colour of the appropriate key was changed from white to yellow. The FI component could be terminated (frustration), and the yellow key darkened, at any time following its initiation. During the time that the FI component was in operation the VI tape ceased to program "reinforcements" to the VI component, and baseline periods failed to be primed by the 2-min. timer. The session length was changed from 100 to 180 min., and magazine exposure time was changed from 4½ to 3 sec.

At the end of Experiment 2, the limited hold contingency was removed from the VI component, and the three birds in that experiment were given 12 sessions adaptation to the new schedule. On a random 50 per cent of occasions the FI was terminated 34 sec. before reward became available, and the VI schedule was again set in operation without food reward being given. On the other 50 per cent of occasions reward was available after 36 sec. Before each of the next 17 sessions began the termination point of the FI was moved 2 sec. closer to the completion of the full (36 sec.) FI. Hence, during the seventeenth session frustration occurred at the completion of the FI component. Following this, the point of FI termination was shifted back 34 sec. for three sessions. The previous procedure was then repeated, the point of FI termination again being moved 2 sec. closer to its full completion time before the start of each session. The replication was complete after 17 sessions.

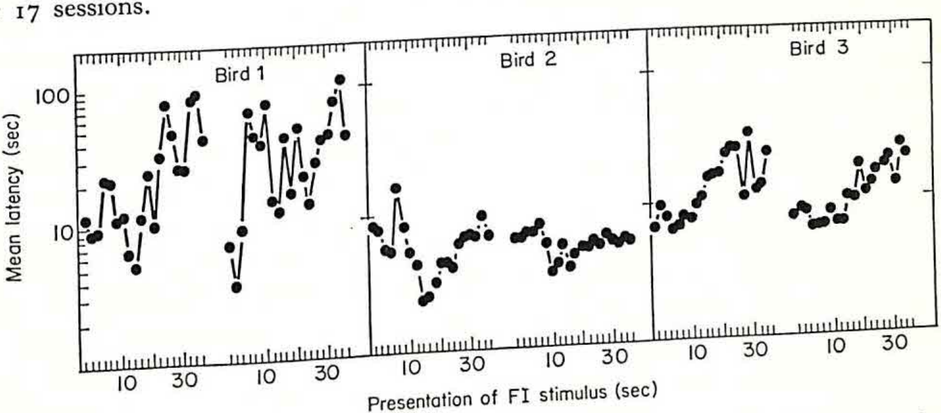


FIGURE 6. Mean sessional time taken to return to variable interval responding following the frustrative termination of the fixed interval in Experiment 3. Latencies are plotted as a function of the time at which frustrative termination occurs. The first three points, in both the first run and replication, are taken when the FI terminates 2 sec. after its onset. The time of each successive termination is then 2 sec. longer than the previous termination. Final termination time (last point) is 36 sec. Note log scale on ordinate.

Results

At no stage in the experiment were there consistent differences between post-frustration and baseline VI rates. Nor did either rate show a consistent change as the termination point of FI neared the completion of FI.

Figure 6 shows the time taken to resume responding on the VI schedule following the termination of the incompleted fixed interval. In the first experimental run, the data from birds 1 and 3 form a monotonic function; latency consistently increases as the time of termination of the fixed interval nears the full completion time of 36 seconds. Bird 2 shows a U-shaped function; latency of return to the VI schedule first decreases and then increases as the termination time of the FI nears the time of completion. In the replication study, the first run functions from all birds appear to be replicated. It should be noted, however, that the first three points indicated for each bird, on both the first and the replication run, are points taken over three initial sessions at the same termination time of 2 sec. Because the monotonic function in bird 1 is so much dependent on the data from these first three sessions, taken at the same termination time, some caution must be exercised in accepting these data from bird 1 as providing an adequate replication.

Discussion

In Experiment 1, where rats were used as subjects, frustration was followed by response energizing effects which increased with increasing magnitudes of reward omitted in the frustrating situation. However, when frustration was produced by the omission of the largest magnitude of reward used in the study, response attenuation sometimes occurred. In Experiments 2 and 3, where pigeons were used as subjects, results taken on the basis of a post-frustration response measure analogous to that used in Experiment 1, failed to duplicate the results of Experiment 1, and no post-frustration increase or decrease in response latency was observed with this measure. But both an increase and a decrease in response latency was observed in Experiment 3. Two birds showed a consistent increase in latency as a frustrative termination of an FI occurred nearer to the completion of that FI. The third bird showed first a decrease in latency and then an increase in latency as the termination of the FI occurred nearer to the completion of the FI, and this function replicated.

The general picture that emerges from the three experiments is that frustration has energizing or depression effects according to its intensity, which presumably varies between animals, but which increases with reward magnitude, and as it occurs nearer to the time of reinforcement. Furthermore, it seems likely that a response depression is caused by competing emotional behaviour, which may take the form of general agitation (Miller and Stevenson, 1936), urination (Amsel, 1958) or pacing (Azrin *et al.*, 1966).

The present results have relevance to frustration theory (Amsel, 1958), one of the most prominent of current theories of partial reinforcement and extinction effects. A basic assumption in frustration theory is that some of the effects of non-reward can be accounted for by the energizing and aversive properties of the frustration response (RF) which it supposedly generates. Most of the direct experimental tests of the basic assumptions have employed an apparatus consisting of two runways placed in series. In this double runway apparatus, an energizing effect of frustration, produced by an omission of reward at the end of the first runway, can be demonstrated as an elevation of running speed in the second runway (e.g. Wagner, 1959). This energizing effect, the difference between post-frustration and baseline running speeds, is called the frustration effect (FE).

An important implication of the present results seems to be that the magnitude of the theoretical construct RF cannot be easily inferred from the magnitude of the FE. As frustration can be followed by response depression, it seems probable that the manipulation of an independent variable which might be expected to increase the magnitude of RF, could be followed by a reduction in the magnitude of the FE, or by no change in the FE, even though other effects of frustration, such as emotional or escape behaviour, might be increasing in strength.

This conclusion should, however, be tempered by a reservation. The FE is measured as a change in the vigour of a "tightly bound" running response which often occurs in a very narrow runway. Such a response may be less susceptible to emotional disruption than those recorded in the present study.

The preceding discussion has particular relevance to those studies which have investigated the magnitude of the FE as a function of reward magnitude in a double runway situation. Peckham and Amsel (1964) ran a single group of rats on a discrimination procedure whereby they were rewarded with 8 pellets of food at the end of the first runway when it was (say) black, and with 2 pellets when it was white. When food was omitted from the end of the first runway (frustration), running speed in the second runway was faster when the first runway (colour) had been associated with an 8 pellet, rather than a 2 pellet reward. Thus the FE in this study was greater, the greater the reduction in reward magnitude. McHose and Ludvigson (1965), and Barrett, Peyser and McHose (1965) trained different groups of rats on different reward magnitudes in the first runway. When reward was omitted from this first runway, running speed in the second runway was no greater for those animals which had previously experienced the largest reward in the first runway than those animals which had previously received a smaller reward in the first runway. One possible explanation of the apparent discrepancy in results between the studies is that the magnitude of the frustration response (and running speed in the second runway) is differentially related to reward magnitude only when the *same* rats experience different reward magnitudes. Another explanation, consistent with the response depression obtained in the present studies, is that the magnitude of the frustration response (RF) caused by the omission of a large reward is so great as to occasionally disrupt, rather than enhance subsequent responding. In any group of animals experiencing the omission of a large reward, this disruption might occur always for some animals or intermittently for all animals. In either case an increase in the magnitude of the frustration response might not then be reflected in an increased measure of its energizing effect. It is worth noting that rats reduced from either 12 to 0, or from 8 to 0 pellets in the Barrett *et al.* study actually ran more slowly in the second runway than rats reduced from 4 to 0 pellets.

References

AMSEL, A. (1958). The role of frustrative non-reward in non-continuous reward situations. *Psychol. Bull.* **55**, 102–19.

AZRIN, N. H., HUTCHINSON, R. R. and HAKE, D. F. (1966). Extinction-induced aggression. *J. exp. Anal. Behav.* **9**, 191–204.

BARRETT, R. J., PEYSER, C. S. and MCHOSE, J. H. (1965). Effects of complete and incomplete reward reduction on a subsequent response. *Psychon. Sci.* **3**, 277–8.

FERSTER, C. B. and SKINNER, B. F. (1957). *Schedules of Reinforcement.* New York: Appleton–Century–Crofts.

FLESHLER, M. and HOFFMAN, H. S. (1962). A progression for generating variable interval schedules. *J. exp. Anal. Behav.* **5**, 529–30.

HAMM, H. D. (1967). Perseveration and summation of the frustration effect. *J. exp. Psychol.* **73**, 196–204.

KELLER, F. S. and SCHOENFELD, W. N. (1950). *Principles of Psychology.* New York: Appleton–Century–Crofts.

MALTZMAN, I. (1952). The process need. *Psychol. Rev.* **59**, 40–8.

MCHOSE, J. and LUDVIGSON, H. (1965). The role of reward magnitude and incomplete reduction in reward magnitude in the frustration effect. *J. exp. Psychol.* **70**, 490–5.

MILLER, N. E. and STEVENSON, S. S. (1936). Agitated behaviour of rats during experimental extinction and a curve of spontaneous recovery. *J. comp. Psychol.* **21**, 205–31.

PECKHAM, R. and AMSEL, A. (1964). Magnitude of reward and the frustration effect in a within-subjects design. *Psychon. Sci.* **1**, 285–6.

SKINNER, B. F. (1938). *The behaviour of organisms.* New York: Appleton–Century–Crofts.

SKINNER, B. F. (1950). Are theories of learning necessary? *Psychol. Rev.* **57**, 193–216.

STADDON, J. E. R. and INNIS, N. K. (1966). An effort analogous to "frustration" on interval reinforcement schedules. *Psychon. Sci.* **4**, 287–8.

TERRACE, H. S. (1966). Stimulus control. In W. K. HONIG (ed.), *Operant behaviour: Areas of Research and Application*, pp. 271–344. New York: Appleton–Century–Crofts.

THOMPSON, T., HEISTAD, G. T. and PALERMO, D. S. (1963). Effect of amount of training on rate and duration of responding during extinction. *J. exp. Anal. Behav.* **6**, 155–61.

WAGNER, A. R. (1959). The role of reinforcement and nonreinforcement in an "apparent frustration effect." *J. exp. Psychol.* **57**, 130–6.

Received 19 *August* 1968

Q. Jl exp. Psychol. (1969) **21**, 381–383

READING AND TYPING

L. H. SHAFFER AND J. HARDWICK
Department of Psychology, University of Exeter

As an extension to an earlier study on typing, an experiment was carried out on skilled touch typists in which they were asked to type texts that differed in word structure and word length. It was predicted from certain assumptions about the reading process that there should be significant effects of structure, length and their interaction. These predictions were confirmed.

Introduction

An earlier paper (Shaffer and Hardwick, 1968) examined the relationship between text structure and the typing speed and accuracy of professional touch typists. In the major experiment the following conjecture was tested: that typing speed is a function of (a) word structure (i.e. the extent to which a word approximates a linguistic form), (b) word length and (c) an interaction of these factors. An analysis of variance on the data led to a rejection of the null hypothesis only for the first of these factors. A *post mortem* on the failure to support (b) and (c) suggested that the experimental test was weak, since word length varied randomly in the texts and only average length differed between texts.

It is proposed here to present a stronger test using a more systematic control of word length in different texts.

Reasons for the conjecture were not given in the earlier paper but in discussing the results it was put forward as hypothesis that typing is controlled by a transcription mechanism that reads linguistic fragments of text as input and produces as output an ordered sequence of response instructions, each instruction covering a single key response. The reasons for the conjecture are stated here as a set of assumptions: (1) less familiar or less structured words are broken into shorter segments, (2) longer words present greater freedom of partitioning into segments and (3) there is a decision time associated with the choice of segments that increases with the number of segments and with the freedom of choice. The postulated interaction arises from the expectation that in (3), choice of number of segments and of segment boundaries are not independent.

Method

Subjects

These were eight touch typists employed in the University. All had served in the earlier experiments and were selected for their accuracy.

Apparatus

Each girl used her own typewriter. Texts were printed completely in lower case 10-point type with 1½ spacing between lines. Times per text were recorded to 0·1 sec. with a Venner digital timer.

Text

The eight texts were as follows: *Real Word:* one *text* of 52 five-letter words and one of 17 words between 15 and 17 letters long. All words were taken from the range 1 : 1,000,000 to 1 : 4,000,000 of the Thorndike-Lorge list. *Syllabic:* one text constructed from each of the Real Word texts by shuffling letters within a word to make a nonsense word having syllabic structure. *First-Order:* one text of 5-letter words and one of 15-letter words, the words constructed randomly with the letter probabilities of written English. *Zero-Order:* one text of 5-letter words and one of 15-letter words, the words constructed randomly with all letters equally likely to occur, but rearranged so as to reduce accidental syllabic structure to a minimum.

Design

Each typist typed each of the texts in a random order based upon a latin square.

Instruction

This is given in detail in the earlier paper. The typists were asked to transcribe the texts as accurately and as quickly as possible.

Results

Only time scores are examined because error rate was less than 1 per cent. Times per test were converted to mean time per symbol, making allowances for carriage return time (see earlier paper).

The average time per symbol for all subjects is shown in Table I. With one exception, Real Word texts, long words take longer to type per symbol than short words and the difference increases as linguistic structure is removed from the words. An analysis of variance on the data shows that type of word is a significant variable, $F\,(3,42) = 208{\cdot}0$, $P < 0{\cdot}001$, so is length of word, $F\,(1,42) = 19{\cdot}25$, $P < 0{\cdot}001$; and so is the interaction of type and length, $F\,(3,42) = 6{\cdot}75$, $P < 0{\cdot}001$.

TABLE I

Average response times per symbol (msec.)

	Real Word	Syllabic	First Order	Zero Order
Long words	178	219	274	371
Short words	190	202	236	325
Difference	−12	17	38	46

The first of the significant results replicates those obtained before; the second is partly new because in the earlier experiments it was only tested and found to hold with Zero Order texts; the third is wholly new and shows that the combined effect exceeds the sum of the individual main effects. This last result, however, is not quite as convincing as it might seem. If one uses the Scheffé S-method of contrasts then it can be shown that the interaction effect depends to a large extent on the reversal of the word length effect in Real Word texts.

Discussion

The results confirm the predictions made above, except for the anomalous result in Real Word texts. The anomaly is easily explained and is not of great interest: in our choice of words having comparable frequencies of usage the short words were not only unfamiliar, they often had unusual syllabic structure (e.g. delft, tinct, aorta), whereas the long words were often concatenations of more familiar words with prefixes and suffixes (e.g. circumnavigation, electromagnetism).

It was stated at the outset that we assume typing speed is limited by the rate at which text can be read. It should be made clear that the term "read" is used generically to refer to a family of input processes each governed by particular response requirements. Since the response requirement in typing is a letter by letter transliteration of text the reading operation must partition the text into segments that can be decomposed into letter strings with some acceptably low risk of loss or error through memory failure. It would thus be misleading to say that people can read faster than they can type. On the other hand, it would be interesting to know whether a typist can adjust her eye-finger span (cf. Butsch, 1932) independently of her typing speed, or whether she can simultaneously read for typing and read for comprehension.

We gratefully acknowledge the financial support of the Science Research Council for this research, which is part of a study on serial skills.

References

BUTSCH, R. L. C. (1932). Eye movements and the eye-hand span in typewriting. *J. educ. Psychol.* **23**, 104–21.

SHAFFER, L. H. and HARDWICK, J. (1968). Typing performance as a function of text. *Q. Jl exp. Psychol.* **20**, 360–9.

Received 21 *October* 1969

Q. Jl exp. Psychol. (1969) **21**, 384–386

BOOK REVIEWS

Techniques of Photostimulation in Biology. Edited by B. H. Crawford. Amsterdam: North-Holland Publishing Company. 1968. Pp. 317. £5 17s.

In spite of the title, this book is biased quite strongly towards the type of visual science that makes use of Maxwellian-view optical systems and the like. Non-visual photobiologists will find little specific help here, nor will more "perceptual" psychologists. The number of topics which are omitted or mentioned only in passing is disappointingly large. For example, the chapter on "Control of the time pattern of the stimulus" does not mention glow-modulator tubes, the use of a CRO as a light source, electroluminescent panels, nor even the standard tachistoscopic electronics-plus-fluorescent-tube apparatus. That on "Control of spectral composition and intensity of stimulus" deals mainly with interference filters and monochromators and says very little about the commoner problems of intensity control. However, there are certainly useful tips here for anyone starting research in the above-mentioned field, and for such people the book is a useful supplement to (but no substitute for) visits to one or two laboratories where such research is being done.

PAUL WHITTLE

The Figure-ground Phenomenon in Experimental and Phenomenological Psychology. By Jan C. Bouman. Stockholm: published by the author. 1968. Pp. 269.

This provides quite a useful international summary of work on the figure-ground division since Rubin. However, it is unlikely that Bouman will convert many readers of this journal to his main thesis, that experimental psychology is in need of revitalization by an injection of existential phenomenology.

PAUL WHITTLE

Methods for the Measurement of Psychological Performance. (IBP Handbook No. 10). By S. Bieshevel. Oxford and Edinburgh: Blackwell Scientific Publications, 1969. Pp. iv + 98. 20s.

This is a handbook designed for the use of workers wishing to participate in the International Biological Programme. The basic notion behind the book is that "man's adaptation to environmental change takes place mainly by way of learning and through cultural means," and the text sets out to describe psychophysiological and human performance data which may be of use in cross-cultural comparisons. The tests described are a rather selected sample, and whilst the choice might be of use within the context of the I.B.P., for general use such a restricted set of tests sharply limits the value of the book.

There are six chapters, covering the acquisition of psychophysiological, performance, perceptual and higher mental data, with some tests described in detail. Many pages are allocated to demonstrations of what general data to log, how to lay data out, etc.

However, I cannot but feel that a scientist proposing to measure any variable described will already have higher-level skills than the book would provide. If he did not, then he would find insufficient detail and have to refer to other sources.

As a teaching book the restricted set of measures described and the total absence of redundancy in the more important paragraphs makes it rather weak. A book such as Joseph Sidowski's "Experimental Methods and Instrumentation in Psychology" or Peter Venables' "Manual of Psycho-physiological methods" would be vastly better (albeit more expensive) aids to teaching techniques of research. The only thing these two lack which are described in the present book is an excellent description of the "Modernity Scale", and some mention of "mental testing" in general.

A. J. SANFORD

Experimental Neuropsychology—A Laboratory Manual. By Benjamin L. Hart. San Francisco: W. H. Freeman and Company. 1969. Pp. 102. £2.

This short text is designed to accompany a practical course of about 40 hours at a late undergraduate or early postgraduate level. There are three sections: the dissection, staining, and microscopic examination of rat and ruminant brains; the techniques of stereotaxic surgery, anaesthesia and perfusion; and four experiments in which behaviour is modified by brain lessions or inter-cranial stimulation. The work in the last two sections is entirely on the rat, and the inclusion of a simple but adequate stereotaxic atlas of the rat brain makes the book almost self-contained. It is a pity though that electrical recording techniques are not included. Otherwise, the material is well selected, and the different techniques are clearly explained by a concise text and a large number of drawings, many in strip cartoon form. The absence of an index is annoying, but not disastrous in such a brief and well organized book.

The one important reservation that must be made before recommending this book concerns the inclusion of chronic lesion and implantation experiments. Apart from the very doubtful ethics of allowing unskilled people to perform operations that would cause extreme suffering if they were done badly, it is unlikely that British students would legally be allowed to do these experiments under classroom conditions. In any case, most behavioural techniques can be taught using normal animals, and Hart's book remains a very useful teaching aid.

RICHARD LATTO

Studies in Cognitive Development. Edited by D. Elkind and J. H. Flavell. London: Oxford University Press. 1969. Pp. xix + 502. 84s.

There has recently been a growing interest in Piaget's work and the present volume is the third published in honour of Jean Piaget in the last few years, after Flavell's monograph *The Developmental Psychology of Jean Piaget* (1963) and *Studies in Cognitive Growth* (Bruner, Olver and Greenfield, 1966).

This collection includes 14 papers by Piaget's students and followers and it is divided into three parts *Theoretical perspectives*, *Research perspectives* and *Applied perspectives* dealing in detail with his ideas. Until now most researchers tried either to support or to disprove Piaget's results by repeating his experiments in different problem situations and different conditions, e.g. with groups varying in age, intelligence, cultural and social background, and in the second part of this book we find some interesting experiments of this type. It seems that the most important in this section, Charlesworth's paper, is an attempt to determine the role of "surprise" in cognitive development. This is commendable as in the studies of surprise, main attention has been and still is devoted to motivation and emotions. (There are some exceptions, e.g. Pavlov's orienting reflex.)

The main contribution of this book lies in the first part which imposes original thoughts in the interpretation of the Piagetian system. It is a theoretcial attempt to analyze the formal aspects of his work and the main concepts each as *conservation* and *stage*, and to compare these concepts with investigations into cognitions. Studies in conservation are the most popular topic for Piagetian experiments: in the work under discussion however, Elkind, aims at a detailed conceptual comparison between Piagetian concept formation and what has been studied in experimental psychology as "*concept formation*" (which is in fact *concept attainment*). Indeed, Elkind is right to state that the Piagetian "version of concepts complements rather than supplants the classical one".

While most studies devoted to psychology of thinking have dealt with the *process*, Flavell and Wohlwill analyze in their paper the *form* and *structure* of Piaget-type systems. The authors follow the recent progress in psycho-linguistics. They had the inspired idea to base their own "general model for analysis of the formation of stages of intelligence" on Chomsky's distinction between competence and performance. Thus, the structures, rules and mental operations are distinguished from the actual process of problem solving. This view has been basic to progress in simulation of thinking and heuristic processes as well.

It has been repeatedly emphasized (e.g. by Duncan), that studies in thinking are still without a coherent theory and in my view the emphasis on structural analysis of stages, problem solving situations and modelling of cognition could be very useful way to progress in psychology of thinking; from this book Piaget's logical system appears a good starting point.

IVANA MARKOVA

Publications Received

Books

Purposive Systems. Edited by H. von Foerster, J. D. White, L. J. Peterson, and J. K. Russell. New York: Spartan Books, London: Macmillan. 1969. Pp. xx + 179. £5. 2nd Ed.

Tests of Musical Ability and Appreciation. By H. J. Wing. Cambridge University Press. 1968. Pp. v + 98. 25s.

Experiment at Berkeley. By Joseph Tussman. New York and London: Oxford University Press, 1969. Pp. xiii + 139. 42s. Paperback 15s.

Mathematics for Introductory Statistics. By A. R. Baggaley. New York: John Wiley. 1969. Pp. xiv + 171. 66s. Paperback 27s.

Attitude Change. By C. A. Kiesler, B. E. Collins and N. Miller. New York: John Wiley. 1969. Pp. xii + 386. 79s.

Psychiatric Aftercare. By M. Silverstein. Philadelphia: University of Pennsylvania Press 1969. Pp. viii + 158. 48s.

Task Interruption. By A. van Bergen. Amsterdam: North Holland. 1968. Pp. viii + 352. £5 5s.

Interaction Concepts of Personality. By R. C. Carson. Chicago: Aldine Publishing Co. 1969. Pp. xiv + 336. $7.50.

Paperbacks

Le temps psychologique. By J. C. Levy. Paris: Dunod. 1969. Pp. ix + 106. Fr. 5·00.

Programmed Learning Research. Edited by A. De Brisson. Paris: Dunod. 1969. Pp. xv + 360. Fr. 78·00.

Leadership. Edited by C. A. Gibb. Harmondsworth: Penguin. 1969. Pp. 439. 9s.

Personal Relationships in Psychological Disorders. By Gordon R. Lowe. Harmondsworth, Penguin. 1969. Pp. 143. 6s.

Henderson and Gillespie's Textbook of Psychiatry. Revised by I. R. C. Batchelor. London: Oxford University Press. 1969. Pp. xii + 578. 42s. Boards 60s.

The Nervous System. By P. Nathan. Harmondsworth, Penguin. 1969. Pp. 391. 10s.

INDEX TO AUTHORS

Abrams, K., with Bever, T. G. 280
Anderson, N. S. 248
Atkinson, R. C., with Freund, R. D. and Brelsford, J. W. Jr. 214
Baddeley, A. D., with Gilson, E. Q. 180
Bever, T. G., with Abrams, K. 280
Biederman, G. R. 332
Blackman, D. E., with Hughes, R. N. 55
Bowman, R., with Sutherland, N. S. 69, 330
Brelsford, J. W., Jr., with Freund, R. D. and Atkinson, R. C. 214
Briggs, G. E., with Swanson, J. M. 185
Brown, H. V., with Miles, T. R. 134
Buffardi, L., with Handel, S. 256
Burns, B. D., with Mandl, G. and Pritchard, R. and Webb, C. 299
Cohen, G. 272
Combs, W. C., with Doty, B. A. 58
Coughlin, P. J., with Kinsbourne, M. 67
Deutsch, J. A., with Huppert, F. A. 267
Dodwell, P. C., with Standing, L. G. and Thio, H. 127
Doty, B. A., with Combs, W. C. 58
Dutch, J. 63
Fisher, G. H. 356
Freund, R. D., with Brelsford, J. W. and Atkinson, R. C. 214
Gerver, D. 203
Gilson, E. Q., with Baddeley, A. D. 180
Gogel, W. C. 239
Haber, R. N., with Standing, L. G. 43
Handel, S., with Buffardi, L. 256
Hardwick, J., with Shaffer, L. H. 209, 381
Hinde, R. A. 105
Hughes, R. N., with Blackman, D. E. 55
Humphrey, N. K., with Weiskrantz, L. 225
Huppert, F. A., with Deutsch, J. A. 267
Johnson-Laird, P. N. 1
Johnson-Laird, P. N., with Wason, P. C. 14
Kinsbourne, M., with Coughlin, P. J. 67
Mandl, G., with Burns, B. D. and Pritchard, R. and Webb, C. 299
Markova, I. 29
Miles, T. R., with Brown, H. V. 134
Nerenberg, A., with Wilton, R. N. and Strongman, K. T. 367
Norman, D. A. 85
Oatley, K., with Robertson, A. and Scanlan, P. M. 172
Oatley, K., with Tonge, D. A. 162

Pritchard, R., with Burns, B. D. and Mandl, G. and Webb, C. .. 299
Rashotte, M. E., with Surridge, C. T. .. 156
Robertson, A., with Oatley, D. and Scanlan, P. M. .. 172
Ross, H. E. .. 346
Ryan, J. .. 137
Ryan, J. .. 148
Sampson, H. .. 39
Sanders, A. F., with Schroots, J. J. F. .. 21
Scanlan, P. M., with Oatley, K. and Robertson, A. .. 172
Schroots, J. J. F., with Sanders, A. F. .. 21
Seymour, P. H. K. .. 312
Shaffer, L. H., with Hardwick, J. .. 209, 381
Standing, L. G., with Dodwell, P. C. and Thio, H. .. 127
Standing, L. G., with Haber, R. N. .. 43
Still, A. W. .. 340
Strongman, K. T., with Wilton, R. N. and Nerenberg, A. .. 367
Surridge, C. T., with Rashotte, M. E. .. 156
Sutherland, N. S., with Bowman, R. .. 69, 330
Sutherland, N. S., with Williams, C. .. 77
Swanson, J. M., with Briggs, G. E. .. 185
Thio, H., with Dodwell, P. C. and Standing, L. G. .. 127
Tonge, D. A., with Oatley, K. .. 162
Wason, P. C., with Johnson-Laird, P. N. .. 14
Weiskrantz, L., with Humphrey, N. K. .. 225
Webb, C., with Burns, B. D., Mandl, G. and Pritchard, R. .. 299
Williams, C., with Sutherland, N. S. .. 77
Wilton, R. N., with Strongman, K. T. and Nerenberg, A. .. 367

The Drug Scene in Great Britain

'Journey into Loneliness'

Max M. Glatt, D. J. Pittman, D. G. Gillespie *and* **D. R. Hills**

15s. net paper 28s. net boards

'This book should be required reading for all who are not specialists in the subject themselves. Everyone, from addicts to their most violent denigrators, will find it informative and if its atmosphere of calm, constructive discussion could be spread throughout society, solutions might be more easily found.'

Medical News

Edward Arnold
41 Maddox Street, London, W.1

Reward and Punishment in Human Learning

Elements of a Behavior Theory

By **Joseph Nuttin**
Department of Psychology
University of Louvain
Louvain, Belgium

In collaboration with

Anthony G. Greenwald
Department of Psychology
Ohio State University
Columbus, Ohio

A psychological interpretation of the operation of reward and punishment in the shaping of memory and behavior. The conclusions arrived at through the authors' numerous experiments are strikingly different from those obtained by Thorndike and others working in this field. The research focuses particularly on the role of the learner's task orientation when receiving reward or punishment, the effects of varying the ratio of reward to punishment in the structure of a task, and the determinants of memory for reward and punishment. Discussions of methodological problems associated with research on reward and punishment are also presented.

1968, 205 pp., 79s. 6d.

Academic Press
NEW YORK and **LONDON**
111 Fifth Avenue, New York, N.Y. 10003
Berkeley Square House, London,
W1X 6BA.

Negro and White Children

A PSYCHOLOGICAL STUDY
IN THE RURAL SOUTH

By **E. Earl Baughman**
W. Grant Dahlstrom
Department of Psychology,
University of North Carolina
Chapel Hill, North Carolina

1968, 572 pp., 140s.

Reports the scientific conclusions of a unique four year study of racially-segregated students in a rural economically deprived area of North Carolina.

CONTENTS:

SETTING AND METHODS
Life in Millfield
Procedures Used to Study Millfield's Children

INTELLIGENCE AND ACADEMIC ACHIEVEMENT
Ability Differences between Racial, Sex, and Age Groups
Academic Achievement Differences between Racial, Sex, Age, and Ability Groups
The Relationship of Selected Family Variables to Ability and Academic Achievement

INTELLECTIVE CHANGE
Changes in IQ Among Young Negro Children
Changes in IQ Among White Adolescents
The Effect of Kindergarten Training upon the Intellective Functioning of Negro and White Children.
A Pilot Project To Accelerate the Intellective Development of Four-Year-Old Negro and White Children

PERSONAL ATTRIBUTES
Millfield's Teachers View Their Pupils
Eighth-Grade Children as Assessed by the Minnesota Multiphasic Personality Inventory
Fantasy Behaviour: Achievement Motivation
Fantasy Behaviour: Other Attributes
Correlates of Peer Popularity

PERSONAL PERSPECTIVES
Mothers of Kindergarten-Age Children Report on Their Children, Training Practices, and Families
Eighth-Grade Children Report on Their Worlds

INTEGRATION AND RECOMMENDATIONS
Some Summary Comments and Observations
Narrowing the Gap

REFERENCES.

APPENDIX

AUTHOR INDEX—SUBJECT INDEX.

Academic Press
NEW YORK AND **LONDON**
111 FIFTH AVENUE, NEW YORK, N.Y. 10003
BERKELEY SQUARE HOUSE, LONDON W1X 6BA